LOMONOSOV MOSCOW STATE UNIVERSITY

on the rights of the manuscript

Nikita Igorevich Tsyganov

Optimization of tree-like transport systems of energy markets

Scientific specialty 1.2.3. Theoretical computer science, cybernetics

Dissertation for the degree of
Candidate of Physical and Mathematical Sciences

Translation from Russian

Scientific supervisor:
Doctor of Physics and Mathematics,
Professor
Alexander Alekseevich Vasin

Moscow - 2024

Table of contents

Introduction

This paper is devoted to the study of optimization mathematical models of multi-node energy markets. Energy resources play an important role in people's lives and the global economy. They are used in the electric power industry, industry, transport, heat supply, and construction. These resources are mainly used as an energy source or in the chemical industry, where they are used as feedstock. Plastics, rubber, dyes, medicines, explosives, oils and much more are obtained from hydrocarbons (natural gas, oil, coal).

Usually, energy resources are extracted in places far from consumers. Hydrocarbons, as a rule, lie deep underground, large areas of their occurrence suitable for profitable industrial development are called deposits. To make their extraction possible, the deposits are previously explored and equipped. Wells are drilled to extract natural gas and oil, and quarries (open-pit mining) or mines (closed-pit mining) are used to develop coal deposits.

Hydrocarbons are non-renewable energy sources. Although their accumulations in the bowels of the earth are still huge, and proven reserves alone will last for at least several decades [64–66], much attention is currently being paid to renewable energy sources (RES), such as solar energy, water, wind, waves, tides, and bioenergy. While hydropower has been used on an industrial scale since the end of the XIX century, the popularity of other renewable energy sources has begun to actively grow relatively recently. Over the past decade, solar energy and wind energy have shown significant growth. So, if in 2008 the shares of capacities of these industries in the electric power industry among all renewable energy sources were 1.4% (solar energy) and 10.5% (wind energy), by 2020 they increased to 26.8% and 26.2% respectively [67, 68]. The share of renewable energy sources in electricity generation in 2019 was 26%[1] (in 2020 - 28.4% compared to 61.4% of the share of hydrocarbons [69]), and in total final energy consumption - 11.2% (compared to 8.7% a decade earlier).

The main disadvantage of renewable energy is its strong dependence on climatic, weather and time conditions[2], which results in variability of energy production volumes. This disadvantage can be eliminated by using energy storage devices that smooth out these fluctuations. Despite the rapid growth in the popularity of renewable energy sources, it is not expected that they will be able to catch up with non-renewable hydrocarbon sources in terms of output in the next few decades, however, renewable energy undoubtedly has an advantage due to environmental friendliness and

[1] Excluding traditional biofuels.
[2] This dependence is less relevant to bioenergy and tidal energy.

inexhaustibility.

The third type of energy resources is nuclear fuel used in nuclear power plants (NPP), icebreakers and submarines. Nuclear energy is based on a controlled nuclear chain reaction, accompanied by the release of energy. The key problem of nuclear energy is its safety. The risks of accidents and terrorist attacks, which can lead to man-made and environmental disasters as a result of explosions and releases of radioactive substances into the environment, cause acute disputes regarding the expediency of using nuclear power plants. The largest accidents in the history of nuclear energy are the accidents at the Chernobyl (1986, [5]) and Fukushima-1 (2011, [6]) nuclear power plants. The second problem is the disposal of radioactive waste.

Due to these disadvantages, some countries adhere to a policy of abandoning nuclear energy [70]. Thus, by 2020, Germany has reduced the volume of electricity generation at nuclear power plants by more than half compared to a decade earlier and plans to completely close all operating nuclear reactors in the near future. Energy substitution occurs mainly due to renewable energy sources. However, in general, the global nuclear energy industry continues to develop actively.

Among the advantages of nuclear power plants, it should be noted that there are no emissions of pollutants into the atmosphere (during trouble-free operation), including carbon dioxide, which has a positive effect on reducing the «greenhouse effect». The second advantage is the high energy intensity of the fuel used (uranium and plutonium) compared to hydrocarbons: the complete burning of a kilogram of uranium enriched to 4% provides the same amount of energy released as burning 100 tons of coal [71]. According to Alexey Likhachev, the head of Rosatom, uranium for the operation of the current generation of nuclear power plants will last for 60 years, and in the future, given the new technologies for the reuse of uranium raw materials, for 600 years [72]. In 2020, the share of nuclear energy in electricity generation was 10.2% [69].

The field of application of energy resources is extensive and eventually covers almost all spheres of human activity. Consumers of energy resources are power plants, boiler houses, industrial enterprises, transport companies, the population and many others. Transportation lines are used to deliver energy resources from production sites to consumers. Gas pipelines, oil pipelines, roads and railways, waterways and air routes can act as transportation lines. As a rule, the nodes of energy production and consumption, together with the transportation lines and intermediate nodes connecting them, represent a complex transport network called the energy market.

Often, the producer and consumer of an energy resource are located at a great distance from each other, which makes the share of transportation costs in the final cost of the energy resource significant for the consumer. The construction of new transportation lines or the modernization of existing ones can help reduce transportation costs. However, such transformations almost always require serious capital investments. Therefore, an urgent task for the modern world economy is effective planning of the development of transport systems, in which, taking into account the various set of possible transformations and the economic benefits derived from them, a certain criterion reaches an optimum.

One of these criteria is the public welfare of the energy market - the total utility of energy consumption minus the total costs of its extraction, transportation and development of the transport system. Also, public welfare can be equivalently defined as the total gain (profit) of all market agents - producers, consumers and owners of the transport system. When maximizing public welfare at the optimum point, the greatest «economic effect» is achieved, extracted from the market and then distributed among market agents depending on the nodal prices and in accordance with the rules of market functioning.

The criterion of public welfare is the best possible criterion for a regulated market in the following terms. For any possible «efficient» (or «fair») distribution of gains among market agents, implemented under some scenario of the development of the transport system of the market, the same or better (i.e., the gains of each market agent will not decrease) distribution can be achieved with optimal, in terms of maximizing public welfare, development of the transport system and correct external regulation, in which the achieved optimum is maintained, but if necessary, the gains are redistributed between market agents. Mechanisms such as fixing the nodal prices for energy resources, subsidies, taxes and compensations can be used in regulation.

Also, the criterion of public welfare is closely related to the concept of perfect competition, in which each individual market agent is unable to influence the price, but has complete information about the market and can freely choose partners for transactions. It is known that in conditions of perfect competition in the market, a price is set that balances supply and demand for energy resources, while maximizing public welfare (Arrow K. D. and Debre J., 1954, [7]). However, this statement is true only in the short term, in which the transport system of the market is fixed, and market agents choose only the volumes of energy resources used (in the case of the producer, they are the volumes of production and sales, the consumer - the volumes of purchase and consumption, the owners of transportation lines - the volumes of purchase, transportation and sale) at given nodal prices.

The concepts of social welfare and perfect competition are fundamental in modern economic theory. One of the very first works that laid the foundations of this theory is the classic book by Leon Walras, published in 1874, «Elements of Pure Political Economy, or the theory of social wealth» [8]. It formulates the principle of general economic equilibrium, based on two basic hypotheses - maximum utility and equality of supply and demand. Wald A., Arrow K. D. and

Debre J. prove the existence of this equilibrium [7,9]. Also Debre J. describes the conditions under which the equilibrium is Pareto-optimal in terms of the gains of market agents, and the conditions under which, on the contrary, any Pareto-optimal situation is an equilibrium [10].

In scientific works devoted to the study of multi-node energy markets, models with a fixed transport structure without the possibility of upgrading transportation lines are mainly considered (Crewe M. A., Fernando C. S., Kleindorfer P. R., Davidson M. R., Guessushkina Yu. V., Kraynes E. M., Novikova N. M., Seleznev A. V., Udaltsov Yu. A., Shiryaeva L. V., Hogan V., Vasin A. A., Vasina P. A., Fogelsang I., Edoli E., Fiorenzani S., Vargiolu T., Wu F., Harsha N., Zlotnik A., Sioshansi R., Rudkevich A. M., Roger Z. R.-M., Conrado B.-S., [11–20]). These works mainly study problems associated with the search for competitive equilibrium. Davidson M. R., Guessushkina Yu. V., Kraynes E. M., Novikova N. M., Seleznev A. V., Udaltsov Yu. A., Shiryaeva L. V. describe a mathematical model for optimizing the functioning of the unified energy system of Russia [14]. In this model, the modes of generating equipment loading are selected according to the price requests of generating companies. The following problems that arise when making management decisions are considered: choosing the composition of the included generating equipment, maximizing public welfare of the market for the day ahead, minimizing the costs of generating companies of the balancing market. The first problem is a partially integer optimization problem. The last two are linear programming problems.

Roselon H. considers an optimization model of a multi-node electricity market [21], which consists of producers, consumers and power transmission lines. Producers and consumers are characterized by the functions of production costs and utility, respectively, lines are characterized by the functions of transport losses and capacity. The problem of maximizing public welfare, which is a convex programming problem, is set, and its solution is given. The same paper describes three ways to attract investments for long-term network expansion. Joskou P. L. and Tyrol J. They consider the multi-node electricity market and study the impact on the market power of players in the distribution of rights to transmit electricity through network lines [22]. At the same time, two types of rights are considered: financial and physical.

In the works dealing with the problem of the development of the transport system, the task of minimizing costs at given volumes of consumption and production is mainly considered (Levit B. Yu., Livshits V. N., Gomez P. V., Saraiva J. T., Zhao H.-S., Chen L., Wu T., Choi J., Tran T., Al-Kib A. A., Thomas R., O H. S., Bilinton R., Suleimani K., Mazlum D., Jaber R. A., [23–28]). In the work of Levit B. Yu. and Livshits V. N. the problems of optimal distribution of cargo flows through the transport network and the choice of the most profitable ways of its development in terms of minimizing transportation costs are considered [23], while the nodes of departure and arrival and volumes for each of the transported and interchangeable types of cargo are considered to be set.

Among the works that explore the task of developing transport systems of energy markets in terms of maximizing public welfare, it is worth noting the works [29–32]. In the works of Vasin A. A. and Dylova E. A. [29, 30] a two-node market is considered in conditions of imperfect competition, the task of optimizing capacity is set for it, three possible Nash equilibria are indicated, and the behaviour of the public welfare function is investigated for each type of equilibrium. The same works consider a two-node market in conditions of perfect competition with several transmission lines, as well as a multi-node market in conditions of perfect competition without fixed transport costs, for which some properties are derived. The chain-type market is studied separately in conditions of perfect and imperfect competition. For this type of market, statements are proved that allow, under certain conditions, to determine the flow directions for transmission lines.

In the works of Vasin A. A. and Dolmatova M. S. [31, 32] for a multi-node energy market in which there are variable and fixed transportation costs, the flow structure invariance condition (FSIC) is introduced, in which the flow directions in the lines are constant and do not depend on capacities. A chain-type market with monotonous initial equilibrium prices is being investigated. For such a market, the validity of the FSIC and the supermodularity of the social welfare function over a set of expandable lines are proved, and an algorithm for solving the problem for the case of zero initial capacity is described. For the star-type market, the criterion for the implementation of the FSIC is written out, the property of this transport structure associated with the concepts of supermodularity and submodularity is proved (the properties of supermodularity and submodularity are studied in the works of Khachaturov R. V. and Cherenin V. P. [33–36]), however, algorithms for solving the problem for arbitrary initial capacities are not proposed.

This dissertation work continues the research started in [31, 32] and is devoted to the task of optimizing the transport system of the energy market of one resource in terms of maximizing social welfare. The energy market consists of multiple nodes and many transportation lines connecting them, the transport structure of the market corresponds to a graph of the «tree» type. Each node represents a local market with its own producers and consumers. Producers are characterized by a function of production costs, consumers - by a function of consumption utility. These functions depend on the volume of production and consumption, respectively.

Each transportation line connects two nodes and allows energy resources to move between them in any direction. The line is characterized by the marginal transmission costs for energy resources, the initial capacity (in a particular case, it may be zero, in this case, we assume, the line has not yet been built) and the cost function of increasing capacity (expansion). The latter consists of two components: a variable and a fixed one. The variable component depends on the amount of expansion. Fixed costs do not depend on the volume of expansion and are charged if the line is expanded.

In the case of natural gas or oil markets, fixed costs for the construction of a main pipeline (transportation line) can include the costs of preparing a construction project, renting or buying land, preparing the pipeline route, building facilities for linear pipeline operation, labour costs, fixed costs for the construction of compressor stations and other fixed costs [1].

The variables in the problem under consideration are the production volumes at the nodes and the flows in the transportation lines. At the same time, consumption volumes in nodes are expressed in terms of production volumes and flows, and the capacity of each line is equal to either the initial capacity (if the flow value does not exceed it) or the flow value (if the flow value exceeds the initial capacity). The task is static, i.e. all the initial parameters of the task do not change over time. It is assumed that a certain period of time has been fixed (for example, a year), and all volumes, flows and capacities are considered in relation to this period. Since the costs of expanding the line are charged once, such costs are preliminarily reduced to the same period, taking into account the discount rate, inflation rate and the expected service lifespan of the line.

The key feature of the problem is the consideration of fixed costs when expanding lines, as a result of which the problem moves from the class of convex optimization problems (Karmanov V. G., 1986, [37]) to the class of NP-hard problems (Gary M., Johnson D., 1982, [38]) and requires the development of special solution algorithms. The original problem generalizes two well-known optimization problems. The first is the problem of maximizing the social welfare of the market in conditions of perfect competition (Arrow K. D. and Debre J., 1954, [7]). The second is the transport problem (Kantorovich L. V., Gavurin M. K., Gisevait G. M., Pardalos P. M., [39, 40]).

Another important problem is the development of a way to use the model in practice for planning the development of real energy markets. Although the scope of this model is not limited in any way, it was originally created to describe the functioning of natural gas and oil markets. In this regard, it is necessary to be able to evaluate the initial parameters of the model for such markets.

The **purpose** of this dissertation research is to develop methods for effective planning of the development of transport systems for the energy markets of gas and oil. To achieve this goal, the following **problems** have been set:

1) to describe a model of a multi-node energy market and formulate the problem of optimizing its transport system in terms of maximizing social welfare;
2) to determine the complexity class of the problem;

3) to develop algorithms for solving the problem for various transport structures and to assess their complexity;
4) to develop a method for estimating the initial parameters of the model for the natural gas market, in which consumers do not have access to natural gas, but there is a possibility of their gasification;
5) to apply the developed methods and algorithms to assess the prospects of gasification of any non-gasified region of the Russian Federation.

The **object of the study** is a multi-node energy market with a tree-like transport structure, the **subject of the study** is an optimization model of a multi-node energy market with a tree-like transport structure.

The **scientific novelty** of the research is as follows:

1) the NP-hardness of the considered problem of optimizing the transport system of the energy market is proved;
2) an algorithm for solving an auxiliary problem with a fixed set of expanded lines is developed, and its complexity is estimated for the case of piecewise linear functions;
3) polynomial algorithms for solving the initial problem have been developed for the following special cases:
    - chain-type market with zero initial transmission capacity;
    - chain-type market with monotonous initial equilibrium prices;
4) algorithms for solving the initial problem have been developed in the case of the flow structure invariance condition being met for the following transport structures: «chain», «star», «star-chain»; the average complexity of these algorithms has been studied;
5) mathematical models and methods have been developed for the natural gas market to assess the initial parameters of the problem: transportation cost functions for gas pipelines, production cost functions for gas fields and demand functions for non-gasified nodes; appropriate estimates have been obtained for main and distribution gas pipelines, gas fields and consumers of Irkutsk oblast;
6) an assessment of the prospects for gasification of Irkutsk oblast was carried out, optimal plans for the development of the gas network in terms of maximizing public welfare for various scenarios of taking into account the environmental component were determined.

Theoretical significance. Proven statements and developed algorithms develop the field of mathematical economics, optimization methods and computational methods.

Practical significance. The results of the research and the developed algorithms can be used in planning the development of real gas or oil markets. The optimization model under study, together with the corresponding algorithms, can be adapted for use in other fields (for example, information networks).

The research used **methods** from the following branches of science: optimization, theory of algorithms, computational mathematics, data analysis, mathematical analysis and discrete mathematics. Various software development methods were also used.

The reliability of the theoretical results obtained is due to the rigor of the proofs of the formulated mathematical statements and is confirmed by the conducted computational experiments. The reliability of the results obtained using data from open sources depends on the reliability of the data itself. The author of the study gave preference to official sources when selecting data.

Approbation. The results of the research were reported at the following mathematical conferences and seminars:

- XVII Baikal International School-seminar «Optimization methods and their applications» (Maximikha village, Republic of Buryatia, 2017);
- II All-Russian Conference «Sociophysics and Socioengineering» (Moscow, 2018);
- 5th International Conference on Energy, Sustainability and Climate Change (Greece, 2018);
- IX Moscow International Conference on Operations Research (Moscow, 2018);
- Seminar on mathematical economics (heads: V. Danilov I., Polterovich V. M.) «Energy markets: optimization of transmission networks» (CEMI RAS, Moscow, March 5, 2019);
- 30th European Conference on Operational Research (Dublin, Ireland, 2019);
- X International Conference «Optimization and Applications» (Montenegro, 2019);
- Scientific conference «Tikhonov Readings 2019» (Moscow, 2019);
- Lomonosov Readings 2020. Computational Mathematics and Cybernetics Section (Moscow, 2020);
- IV Russian Economic Congress «REC-2020» (Moscow, 2020);
- Scientific conference «Tikhonov Readings 2021» (Moscow, 2021);
- The 9th International Conference on Information Technology and Quantitative Management (China, 2022);
- Seminar on mathematical economics (heads: V. Danilov I., Polterovich V. M.) «Optimization of transport systems of energy markets» (CEMI RAS, Moscow, November 7, 2023).

Publications. The main results on the research topic have been published in 10 printed publications: 3 - in journals recommended by the Higher Attestation Commission [41, 43, 44], 7 - in collections and abstracts [42, 45–50].

Main scientific results. The following main scientific results were obtained during the study:

1) • for a star-type multi-node energy market of a single resource (and the more general case of the tree-type market), in which fixed costs are present during the expansion of transportation lines that do not depend on the volume of expansion, the task of optimizing the transport system in terms of maximizing social welfare (hereinafter referred to as the *initial problem*) is NP-hard; see [43] (the personal contribution of the author of the dissertation is 100%);

2) • for the auxiliary problem of optimizing a transport system with a fixed set of expanded lines (hereinafter referred to as the *auxiliary problem*), which is a convex programming problem, a special solution algorithm has been developed; see [43] (the personal contribution of the author of the dissertation is 100%);

• its complexity is estimated for the case of piecewise linear initial functions: the number of computational operations of the algorithm does not exceed the value of some predetermined quadratic function of the number of nodes in the market; see [43] (the personal contribution of the author of the dissertation is 100%);

3) • algorithms for solving the initial problem for various transport structures have been developed for the case of the flow structure invariance condition being met, in which the directions of flows in transmission lines are constant and do not depend on transmission capacities; algorithms have been developed for the following transport structures: «chain», «star», «star-chain»; see [41, 42] («chain») [44] («star»); the personal contribution of the author of the dissertation to the direct development of algorithms is 100%; the theorem on the property of complementary and competitive lines underlying the algorithms was formulated and proved by other authors earlier;

• the average statistical complexity of these algorithms is investigated: for computational experiments with random generation of initial problems for each of the three cases, even for a large number of nodes (more than 50), the initial problem is solved in a reasonable time, and the dependence of the average number of solved auxiliary problems on the number of nodes in the market is approximated by a quadratic function; see [41, 42] («chain») [44] («star»); the personal contribution of the author of the dissertation to the implementation and evaluation of algorithms is 100%;

4) • for the natural gas market, methods have been developed for estimating the demand functions for non-gasified nodes; see [45] (the personal contribution of the author of the dissertation to the development of models is at least 80%, the contribution to the derivation of demand functions together with evidence is 100%);

5) • the developed algorithms and methods are applied to assess the prospects of gasification of Irkutsk oblast with the possibility of connecting thermal power plants and boiler houses of the region to the main gas pipeline «Power of Siberia»; see [45] (the personal contribution of the author of the dissertation is 100%);

- according to the calculations based on data on the characteristics of thermal power plants and boiler houses in the region for 2021-2022, gasification brings a positive effect only if the environmental component is taken into account, expressed in the form of a fine for burning each unit of coal currently used in the region; see [45] (the personal contribution of the author of the dissertation is 100%);
- the calculations performed showed that the developed algorithms can be used in planning the development of real energy markets and allow solving the initial problem in a reasonable time.

Provisions to be defended. Based on the results of the study, the following provisions are submitted for defence:

1) • for a star-type multi-node energy market of a single resource (and the more general case of the tree-type market), in which fixed costs are present during the expansion of transportation lines that do not depend on the volume of expansion, the task of optimizing the transport system in terms of maximizing social welfare is NP-hard;

2) • for the auxiliary problem of optimizing a transport system with a fixed set of expanded lines, which is a convex programming problem, a special solution algorithm exists;

- for the case of piecewise linear initial functions the number of computational operations of this algorithm does not exceed the value of some predetermined quadratic function of the number of nodes in the market;

3) • algorithms for solving the initial problem for various transport structures exist for the case of the flow structure invariance condition being met, in which the directions of flows in transmission lines are constant and do not depend on transmission capacities;

- for computational experiments with random generation of initial problems for «chain», «star», and «star-chain» transport structures the dependence of the average number of solved auxiliary problems on the number of nodes in the market is approximated by a quadratic function;

4) • according to the calculations based on data on the characteristics of thermal power plants and boiler houses in the region for 2021-2022, gasification of Irkutsk oblast with the possibility of connecting thermal power plants and boiler houses of the region to the main gas pipeline «Power of Siberia» brings a positive effect only if the environmental

component is taken into account, expressed in the form of a fine for burning each unit of coal currently used in the region;

- the calculations performed showed that the developed algorithms can be used in planning the development of real energy markets and allow solving the initial problem in a reasonable time.

Chapter 1. Algorithms for optimizing the transport system of the energy market

This chapter is devoted to the study of the optimization model of the multi-node energy market. The market consists of many nodes where producers and consumers of energy resources are located, as well as a tree-like transport structure that allows energy resources to move between nodes. The optimization criterion is social welfare - the total utility of consumption minus the total costs of production and transmission.

In order to reduce the recording and to ease perception throughout the chapter, the term *«market»* is used instead of *«energy market»*, instead of *«energy resource»* - *«resource»*, and instead of *«transportation line»* - *«line»*.

Paragraph 1.1. describes the formulation of the problem of optimizing the energy market transport system. The auxiliary problem of maximizing social welfare with a fixed set of expandable lines is formulated. It is shown that the initial problem reduces to the problem of finding the optimal set of expandable lines.

In paragraph 1.2. the NP-hardness of the initial problem is proved.

Paragraph 1.3. is devoted to the study of the auxiliary problem. It introduces the concept of competitive equilibrium, which is closely related to the solution of the auxiliary problem. An algorithm for solving this problem is proposed, and its complexity is estimated for the case of piecewise linear functions.

Paragraph 1.4. introduces the flow structure invariance condition, in which the flow direction is constant for each line and is known in advance. The relations of complementarity and competitiveness for lines are also defined there. It shows how these relations can be used in the search for the optimal set of expandable lines.

Paragraph 1.5. discusses some special cases of the problem for which there are polynomial algorithms for solving, describes the corresponding algorithms and evaluates their complexity.

In paragraph 1.6. algorithms for solving the problem, if the flow structure invariance condition is met, are proposed for the following transport structures: «chain», «star», «star-chain» and «tree». For each algorithm, a computational experiment is conducted to estimate the average statistical complexity, and an approximation of the average number of auxiliary problems to be solved, depending on the number of nodes, is found.

1.1. Formulation of the problem

Let's consider a model of a multi-node market for a homogeneous[3] product consisting of multiple local markets and a transmission network. The functioning of the model is considered in the long term, but it is assumed that the market structure does not change significantly in dynamics. At the same time, a certain basic period of time is allocated (for example, a year), in relation to which the values used in the model and representing the volume of goods are measured (for example, the volume of production per year, the volume of consumption per year, the volume of transmission per year). For the sake of brevity, this gap is not mentioned later in this chapter, but it is implied. The undirected[4] graph $G = (N, L)$ characterizes the transport structure of the market and consists of a set of nodes (local markets) N and a set of lines $L \subseteq \{\{i, j\} \mid i, j \in N\}$. We assume that any pair of nodes is connected by at most one line, and graph G is a tree, i.e. it does not contain cycles (figure 1). The set L can include both existing lines and potential ones that have not yet been built.

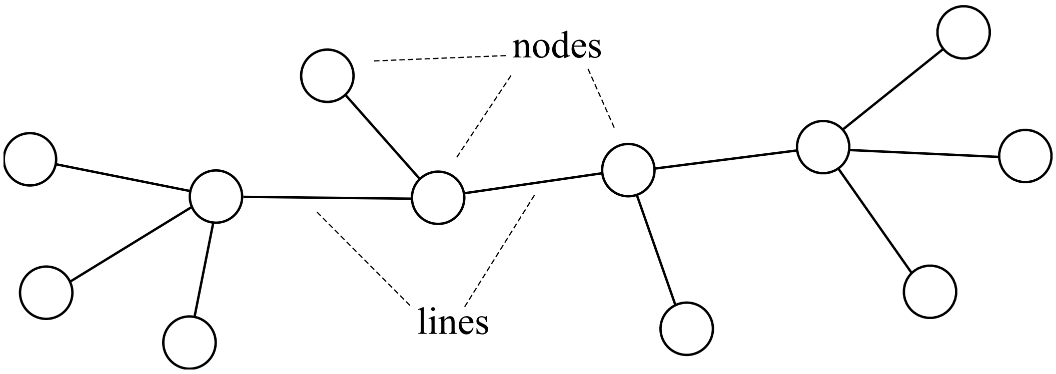

Fig. 1. An example of a tree-type transport market structure

1.1.1. Nodes

Each node $i \in N$ represents a local market where producers and consumers of goods can be present.

The producers of node i are characterized by the <u>*function of production costs*</u> $c_i(v_i)$, depending on the volume of production v_i and defined on the set $[0, V_i^{\max}]$, where the maximum

[3] That is, not differing in quality (Vasin A. A., Morozov V. V., 2005, [51]).

[4] The requirement of graph undirectidness is introduced for the convenience of description. For all the results obtained in the work, their analogues can be obtained for the case of a directed graph.

volume of production $V_i^{\max} \geq 0$ belongs to the extended numerical line and can be equal to $+\infty$, if the volume of production is not limited. In a special case $V_i^{\max}$ can also be zero, which means that there is no possibility of production at this node. Denote for an arbitrary function f by f'_- (f'_+) the left (respectively right) derivative. We assume that the function $c_i(v_i)$ is continuous, non-decreasing, convex, $c_i(0) = 0$ and $c'_{i-}(v_i) \to +\infty$ for $v_i \to +\infty$, if $V_i^{\max} = +\infty$. The economic meaning of these properties is intuitive. With an increase in production, we assume that costs cannot decrease. The convexity property means that enterprises introduce production capacities in ascending order of marginal costs, thus minimizing their own costs. The latter property suggests that marginal costs tend to infinity with increasing output.

Consumers of node i are characterized by the *function of consumption utility* $U_i(v_i^d)$, depending on the volume of consumption $v_i^d \geq 0$. We assume that this function is continuous, does not decrease, is concave, and is equal to a constant for sufficiently large v_i^d, $U_i(0) = 0$ and $U'_{i+}(0) < +\infty$. It is worth explaining the meaning of these assumptions. If the product is not consumed, then the utility is zero, while the marginal utility is limited in zero. With an increase in consumption, the utility can only increase. The concavity property means that the utility of purchasing another unit of goods is no greater than that of purchasing the previous unit. We also assume that there is a certain finite amount of saturation, at which a further increase in consumption no longer leads to an increase in utility. Let's denote it by $V_i^{d,\max}$.

1.1.2. Lines

Each line $\{i, j\} \in L$ allows the product to move between nodes i and j in any[5] direction. We assume that the marginal cost $e^t_{\{i,j\}} \geq 0$ of transmission of goods is constant and does not depend on the direction of flow, and the volume of goods transferred is limited by the initial transmission capacity of the line $Q^0_{\{i,j\}} \geq 0$, which can be increased. The cost of increasing transmission capacity consists of two components: a fixed $E^f_{\{i,j\}} \geq 0$ and a variable $E^v_{\{i,j\}}(\Delta Q_{\{i,j\}})$, depending on the volume of expansion (increase in transmission capacity) of the line.

The presence of a fixed component is due to the fact that during the construction or expansion of the line there are costs, the value of which slightly depends or does not depend at all on the volume of expansion, but is determined only by the fact of expansion. Such costs include the costs of designing, renting land, preparing the route of the line, labour costs and other

[5] In the case of a problem with an oriented graph, movement is possible only in one direction.

fixed costs.

We assume that the variable component $E^v_{\{i,j\}}(\Delta Q_{\{i,j\}})$ is continuous, non-decreasing, convex, defined on the set $[0, \Delta Q^{\max}_{\{i,j\}}]$, where the maximum volume of expansion $\Delta Q^{\max}_{\{i,j\}} \geq 0$ belongs to the extended numeric line and can be equal to $+\infty$, and also that $E^v_{\{i,j\}}(0) = 0$.

It is worth noting that the costs of expanding the line are of a one-time nature, while the production, transfer and consumption of goods are carried out constantly during the functioning of the market. This inconsistency is eliminated by bringing the cost of expanding the line to the base time interval in which the volumes of production, transmission and consumption of goods are considered. Let the base interval be equal to τ years. The reduction is made taking into account the service life of the line $T_{\{i,j\}}$, which is also measured in years, and the coefficient δ - the difference between the interest rate on bank deposits and the inflation rate (continuous discounting is considered). Let $E^{f,r}_{\{i,j\}}$ and $E^{v,r}_{\{i,j\}}(\Delta Q_{\{i,j\}})$ be the initial fixed and variable costs for the extension of the line, respectively, which are charged once at the initial time. Then the given costs, which are further used in the model, are as follows (Stoft S., 2002, [52]): $E^f_{\{i,j\}} = k \cdot E^{f,r}_{\{i,j\}}$, $E^v_{\{i,j\}}(\Delta Q_{\{i,j\}}) = k \cdot E^{v,r}_{\{i,j\}}(\Delta Q_{\{i,j\}})$, where

$$k = \begin{cases} \dfrac{\delta\tau}{1 - e^{-\delta T_{\{i,j\}}}}, & \delta > 0, \\ \dfrac{\tau}{T_{\{i,j\}}}, & \delta = 0. \end{cases} \tag{1.1}$$

Denote by q_{ij} the volume of transfer (flow[6]) of goods from node i to node j. If $q_{ij} < 0$, then we assume that the goods are transferred in the opposite direction, i.e. $q_{ji} = -q_{ij}$. The *transport costs* $E_{ij}(q_{ij})$ for line $\{i, j\}$ are made up of the cost of transferring the goods and the cost of extending the line if the amount of flow exceeds the initial capacity of the line:

$$E_{ij}(q_{ij}) = \begin{cases} e^t_{\{i,j\}}|q_{ij}|, & |q_{ij}| \in [0, Q^0_{\{i,j\}}], \\ e^t_{\{i,j\}}|q_{ij}| + E^f_{\{i,j\}} + E^v_{\{i,j\}}(|q_{ij}| - Q^0_{\{i,j\}}), & |q_{ij}| \in (Q^0_{\{i,j\}}, Q^{\max}_{\{i,j\}}], \end{cases} \tag{1.2}$$

where $Q^{\max}_{\{i,j\}} = Q^0_{\{i,j\}} + \Delta Q^{\max}_{\{i,j\}}$. The function $E_{ij}(q_{ij})$ is defined on the set $|q_{ij}| \leq Q^{\max}_{\{i,j\}}$ and is non-negative, even, and $E_{ij}(0) = 0$. If $E^f_{\{i,j\}} > 0$ and $\Delta Q^{\max}_{\{i,j\}} > 0$, then it is not convex, since it has a gap at $q_{ij} \in \{-Q^0_{\{i,j\}}, Q^0_{\{i,j\}}\}$ (figure 2). It is worth noting that the functions $E_{ij}(q_{ij})$ and $E_{ji}(q_{ji})$ coincide.

[6] It is worth recalling that the flow, as well as the volumes of production and consumption, are measured in relation to some basic period of time (for example, a year).

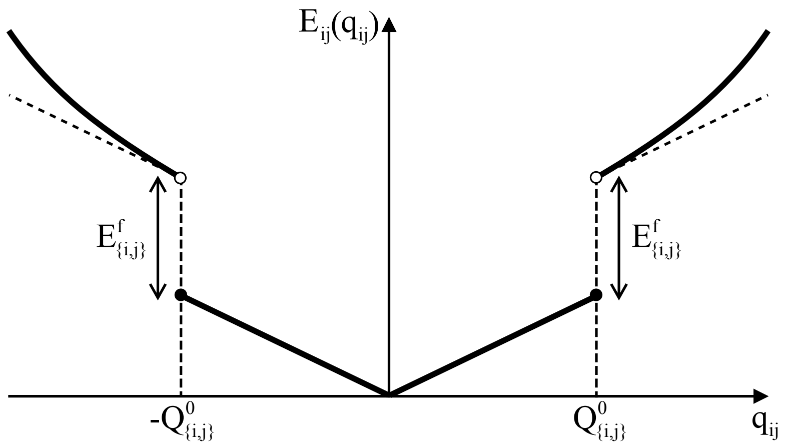

Fig. 2. Example of the transport cost function $E_{ij}(q_{ij})$ with $\Delta Q^{\max}_{\{i,j\}} = +\infty$

1.1.3. Social welfare

Denote by $\vec{q} = (q_{ij} \mid \{i, j\} \in L)$ the flow vector for which the conditions $q_{ji} = -q_{ij},\ \{i, j\} \in L$ are met, by $\vec{v} = (v_i \mid i \in N)$ - the vector of production volumes, and by $Z(i) = \{j \in N \mid \{i, j\} \in L\}$ - the set of adjacent nodes for node $i \in N$. For fixed vectors $\vec{q}$ and $\vec{v}$ we define the consumption volumes $\vec{v^d}(\vec{q}, \vec{v}) = (v_i^d(\vec{q}, \vec{v}) \mid i \in N)$ as follows:

$$v_i^d(\vec{q}, \vec{v}) = v_i - \sum_{j \in Z(i)} q_{ij},\ \ i \in N. \tag{1.3}$$

Here it is assumed that the volume of consumption in the node is equal to the volume of production minus the total outflow of goods from the node. Let's number all nodes from the set N in an arbitrary way and define *social welfare* as the total utility of consumption minus total production and transportation costs:

$$W(\vec{q}, \vec{v}) = \sum_{i \in N} U_i\Big(v_i - \sum_{j \in Z(i)} q_{ij}\Big) - \sum_{i \in N} c_i(v_i) - \sum_{\{i,j\} \in L,\ i<j} E_{ij}(q_{ij}). \tag{1.4}$$

Social welfare can also be represented as the sum of the gains of all market players: producers, consumers and line owners. Denote by $\vec{p} = (p_i \mid i \in N)$ the price vector, where p_i is the price at which the purchase and sale of goods takes place at node i. Then the producers' profit $Pr_i(\vec{q}, \vec{v}, \vec{p})$ and consumer surplus $CS_i(\vec{q}, \vec{v}, \vec{p})$ of node i are equal to the following values:

$$Pr_i(\vec{q}, \vec{v}, \vec{p}) = p_i v_i - c_i(v_i), \qquad CS_i(\vec{q}, \vec{v}, \vec{p}) = U_i\Big(v_i - \sum_{j \in Z(i)} q_{ij}\Big) - p_i\Big(v_i - \sum_{j \in Z(i)} q_{ij}\Big).$$

The profit from the operation of the line $\{i,j\} \in L$ is formed as a result of the purchase of goods in one incident node and its sale in the second node and is equal to

$$T_{ij}(\vec{q},\vec{v},\vec{p}) = (p_j - p_i)q_{ij} - E_{ij}(q_{ij}).$$

The total profit of the transport system is as follows:

$$T(\vec{q},\vec{v},\vec{p}) = \sum_{\{i,j\}\in L,\ i<j} T_{ij}(\vec{q},\vec{v},\vec{p}).$$

As a result, the alternative representation of social welfare has the following form for any price vector $\vec{p}$:

$$\begin{aligned} W(\vec{q},\vec{v}) = &\sum_{i\in N} U_i\Big(v_i - \sum_{j\in Z(i)} q_{ij}\Big) - \sum_{i\in N} c_i(v_i) - \sum_{\{i,j\}\in L,\ i<j} E_{ij}(q_{ij}) \\ &+ \sum_{i\in N}\left(p_i\Big(v_i - \sum_{j\in Z(i)} q_{ij}\Big)\right) - \sum_{i\in N}\left(p_i\Big(v_i - \sum_{j\in Z(i)} q_{ij}\Big)\right) \\ = &\sum_{i\in N}(p_i v_i - c_i(v_i)) + \sum_{i\in N}\left(U_i\Big(v_i - \sum_{j\in Z(i)} q_{ij}\Big) - p_i\Big(v_i - \sum_{j\in Z(i)} q_{ij}\Big)\right) \\ &+ \sum_{\{i,j\}\in L,\ i<j}((p_j - p_i)q_{ij} - E_{ij}(q_{ij})) \\ = &\sum_{i\in N} Pr_i(\vec{q},\vec{v},\vec{p}) + \sum_{i\in N} CS_i(\vec{q},\vec{v},\vec{p}) + T(\vec{q},\vec{v},\vec{p}) \equiv W^a(\vec{q},\vec{v},\vec{p}). \end{aligned}$$

Representations $W(\vec{q},\vec{v})$ and $W^a(\vec{q},\vec{v},\vec{p})$ coincide as a function of $\vec{q},\vec{v}$. Thus, when prices change, there is only a redistribution of social welfare between the players, whereas social welfare itself remains unchanged.

1.1.4. The problem of maximizing social welfare

Let's set the *problem of maximizing social welfare*:

$$W(\vec{q},\vec{v}) \underset{\vec{q},\vec{v}}{\rightarrow} \max. \tag{1.5}$$

The current chapter is devoted to solving the problem (1.5). Its main difficulty lies in the presence of fixed costs $E^f_{\{i,j\}}$ in (1.2), which is why the objective function of the problem is not concave, i.e. the problem is not a convex programming problem (Karmanov V. G., 1986, [37]).

The issue of implementing the optimal state of the market deserves special attention. Let $(\overrightarrow{q^*}, \overrightarrow{v^*})$ is some solution of the problem (1.5). What are the mechanisms of market regulation that stimulate all market players to act in accordance with the found solution: producers to produce goods in volumes $\overrightarrow{v^*}$, the transport system to re-sell them in volumes $\overrightarrow{q^*}$, and consumers to buy in volumes $\vec{v^d}(\overrightarrow{q^*}, \overrightarrow{v^*})$? After all, all players, we assume, are rational and strive to maximize their winnings. As will be shown below, the solution $(\overrightarrow{q^*}, \overrightarrow{v^*})$ corresponds to the vector of equilibrium prices, when fixing which it is beneficial for all market players to act in accordance with the solution $(\overrightarrow{q^*}, \overrightarrow{v^*})$ if the owner of each line $\{i, j\} \in L$ in the case of its expansion receives compensation in the amount of $E^f_{\{i,j\}}$ from a player interested in implementing the optimal state. Such a player may be the state. It is also able to regulate prices in nodes with the help of the Federal Antimonopoly Service (if we are talking about the Russian market), not allowing them to deviate from equilibrium, in the case of the presence of large players in the nodes that can influence prices.

In the case of considering the Russian gas or oil market, the possibility of implementing the optimal state is facilitated by the fact that the largest companies owning main pipelines (in the case of the gas market, «Gazprom», in the case of the oil market, «Transneft»), are controlled by the state.

1.1.5. Reducing the initial problem to the problem of finding the optimal set of extensible lines

Consider an *auxiliary problem* with a fixed set of extensible lines $R \subseteq L$:

$$W(\vec{q}, \vec{v}, R) \underset{\vec{q},\vec{v}}{\rightarrow} \max, \tag{1.6}$$

where $W(\vec{q}, \vec{v}, R)$ is different from $W(\vec{q}, \vec{v})$ in that in the calculation of $W(\vec{q}, \vec{v}, R)$, if $\{i, j\} \in R$, regardless of the flow q_{ij} fixed costs for the extension of the line $E^f_{\{i,j\}}$ are always included in the function of transportation costs (1.2), and if $\{i, j\} \in L \setminus R$, we assume that $Q^{\max}_{\{i,j\}} = Q^0_{\{i,j\}}$ regardless of $\Delta Q^{\max}_{\{i,j\}}$. Let's denote by $\widetilde{W}(R)$ the maximum welfare value in the problem (1.6), and by $\widetilde{X}(R) = \{(\vec{q}, \vec{v}) \mid W(\vec{q}, \vec{v}, R) = \widetilde{W}(R))\}$ - a set of solutions to this problem. Then the original problem (1.5) reduces to the *problem of finding the optimal set of extensible lines*:

$$\widetilde{W}(R) \underset{R\subseteq L}{\to} \max . \tag{1.7}$$

Therefore, if L^* is the solution of the problem (1.7), then $(\overrightarrow{q^*}, \overrightarrow{v^*}) \in \widetilde{X}(L^*)$ is the solution of the problem (1.5).

1.2. NP-hardness of the problem

Let's prove that the original problem (1.5) is NP-hard[7]. Let's consider a special case of the problem in which the market consists of a set of producing nodes $N^s = \{1, ..., n\}$ and a consuming node 0 connected to all the others: $N = N^s \cup \{0\}$, $L = \bigcup\limits_{i\in N^s} \{i, 0\}$. Each producing node $i \in N^s$ is characterized by a maximum production volume $V_i^{\max} = g_i \in \mathbb{N}$, while $c_i(v_i) \equiv 0$, $U_i(v_i^d) \equiv 0$. The consuming node has no production ($V_0^{\max} = 0$) and is described by the utility function of consumption, which is determined by the parameter $K \in \mathbb{N} \cup \{0\}$ and has the following form (figure 3):

$$U_0(v_0^d) = \begin{cases} (K+1)v_0^d - (v_0^d)^2/2, & v_0^d \in [0, K+1], \\ (K+1)^2/2, & v_0^d \in (K+1, +\infty). \end{cases}$$

Each line $\{i, 0\} \in L$ is characterized by fixed extension costs $E^f_{\{i,0\}} = g_i \in \mathbb{N}$, at the same time $e^t_{\{i,0\}} = 0$, $Q^0_{\{i,0\}} = 0$, $\Delta Q^{\max}_{\{i,0\}} = +\infty$, $E^v_{\{i,0\}}(\Delta Q_{\{i,0\}}) \equiv 0$. Then for a fixed set of producing nodes $\overline{N} \subseteq N^s$ with extensible incident lines, the maximum value of social welfare is obviously as follows:

$$\widetilde{W}(\overline{N}) = U_0\Big(\sum_{i\in\overline{N}} V_i^{\max}\Big) - \sum_{i\in\overline{N}} E^f_{\{i,0\}} = U_0\Big(\sum_{i\in\overline{N}} g_i\Big) - \sum_{i\in\overline{N}} g_i.$$

Here, for all expandable lines, the maximum possible volume of goods is transferred to the consuming node. As a result, the problem under consideration is determined by parameters n, K, g_i ($i \in \{1, ..., n\}$) and is equivalent to the following:

$$\widetilde{W}(\overline{N}) \underset{\overline{N}\subseteq N}{\to} \max . \tag{1.8}$$

[7] A problem is called NP-hard if any problem from the NP class is polynomial reduced to it (Gary M., Johnson D., 1982, [38]). The NP class consists of solvability problems that can be solved on a non-deterministic Turing machine in polynomial time of the length of the input data.

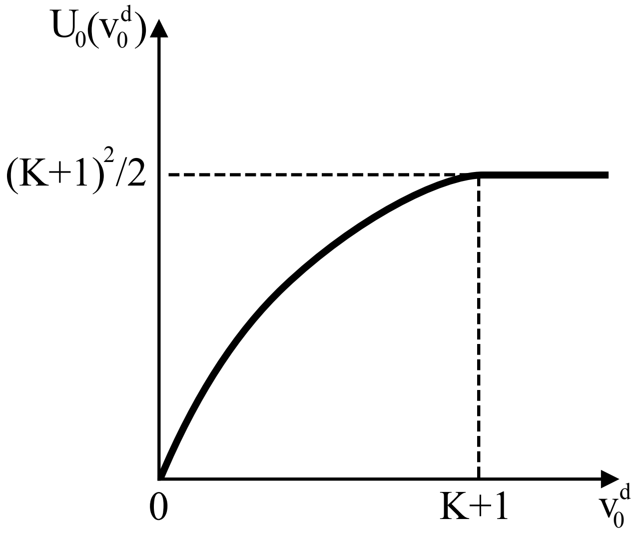

Fig. 3. Consumption utility function $U_0(v_0^d)$

Let 's now consider the following decision problem[8] with an additional parameter $W \in \mathbb{R}$: is there a set

$$\overline{N} \subseteq N, \text{ for which } \widetilde{W}(\overline{N}) \geq W? \tag{1.9}$$

Note that problem (1.9) reduces to problem (1.8) in polynomial time of the length of the input parameters. Indeed, to solve problem (1.9), it is sufficient to find the maximum welfare value W^* in problem (1.8), and then check the inequality $W^* \geq W$. Let's prove the NP-completeness[9] of problem (1.9), which will result in the NP-hardness of problem (1.8) (and hence the more general problem (1.5)).

Theorem 1. *Problem (1.9) is NP-complete.*

Proof. *Problem (1.9) belongs to the NP class, since it can be solved in polynomial time from the length of the input parameters on a non-deterministic Turing machine by considering all subsets of $\overline{N} \subseteq N$ and checking the corresponding inequality for each of them. Therefore, to prove the statement, it is sufficient to reduce to problem (1.9) the well-known NP-complete problem «on the sum of subsets» (Kleinberg D., Tardosh E., 2006, [53]): for a given set of natural numbers $S = \{s_1, ..., s_m\}$ and natural number P is there a subset*

$$S' \subseteq S, \text{ for which } \sum_{s \in S'} s = P? \tag{1.10}$$

Let's consider the problem (1.9) with the following parameters: $n = m$, $K = P$, $g_i = s_i$

[8] The decision problem is a problem with the answer «yes» or «no» (Gary M., Johnson D., 1982, [38]).
[9] The decision problem is called NP-complete if it belongs to the NP class and any other problem from the NP class is polynomial reduced to it (Gary M., Johnson D., 1982, [38]).

($i \in \{1, ..., m\}$), $W = P^2/2$. In that case

$$\widetilde{W}(\overline{N}) = U_0\Big(\sum_{i\in\overline{N}} s_i\Big) - \sum_{i\in\overline{N}} s_i = \varphi\Big(\sum_{i\in\overline{N}} s_i\Big),$$

where

$$\varphi(t) = \begin{cases} Pt - \frac{1}{2}t^2, & t \in [0, P+1], \\ \frac{1}{2}(P+1)^2 - t, & t \in (P+1, +\infty). \end{cases}$$

Note that the function $\varphi(t)$ is concave, and its only maximum is reached at $t = P$, so $\max(\varphi(t)) = \varphi(P) = P^2/2$. This means that the inequality $\widetilde{W}(\overline{N}) \geq P^2/2$ holds if and only if $\sum_{i\in\overline{N}} s_i = P$. Therefore, the answer to the problem (1.9) with these parameters is positive if and only if the answer to the problem (1.10) is positive, i.e. the desired information is obtained. ■

Consequence. *Problem (1.5) is NP-hard.*

The obtained result suggests that, in general, there are no effective[10] methods for solving the problem (1.5). However, there are such methods for some special cases of the problem, which will be shown later.

1.3. Solving the auxiliary problem

1.3.1. Dot-multiple mappings and operations on them

Let's introduce some notations and operations that are used later in the work. For a point-set mapping $G(x)$ defined on the set $X \subseteq \mathbb{R}$, we introduce the following definitions.

Definition 1. *A dot-multiple mapping G is called correct if for any $x \in X$ $G(x)$ is a convex set, and the graph of the mapping $\{(x, g) \mid x \in X,\ g \in G(x)\}$ is a continuous mapping of some one-dimensional convex nonempty set in $\mathbb{R}^2$.*

Definition 2. *A dot-multiple mapping G is called non-decreasing (non-increasing) if for any x_1, $x_2 \in X$, $x_1 < x_2$, and for any $g_1 \in G(x_1)$, $g_2 \in G(x_2)$ the inequality $g_1 \leq g_2$ ($g_1 \geq g_2$) is met.*

Definition 3. *A point-set mapping G is called monotonic if it is non-decreasing or non-increasing.*

[10] That is, solved by a deterministic Turing machine in polynomial time from the length of the input data (Gary M., Johnson D., 1982, [38]).

Let's denote by $P^{\uparrow}$ ($P^{\downarrow}$) the set of all non-decreasing (non-increasing) regular point-multiple maps, and by $P^m = P^{\uparrow} \cup P^{\downarrow}$ the set of all monotonic regular point-multiple maps.

For the sets Y, Y_1, $Y_2 \subseteq \mathbb{R}$, and the number $y \in \mathbb{R}$ we introduce the following operations:

- $-Y \equiv \{y \mid -y \in Y\}$;
- $Y_1 + Y_2 \equiv \{y \mid y = y_1 + y_2,\ y_1 \in Y_1,\ y_2 \in Y_2\}$;
- $Y_1 - Y_2 \equiv Y_1 + (-Y_2)$;
- $Y + y \equiv y + Y \equiv Y + \{y\}$;
- $Y - y \equiv -y + Y \equiv Y - \{y\}$;
- $-Y + y \equiv y - Y \equiv (-Y) + \{y\}$;
- $-Y - y \equiv -y - Y \equiv (-Y) - \{y\}$;
- $\max(Y) \equiv \max\limits_{y \in Y} y$, $\min(Y) \equiv \min\limits_{y \in Y} y$ (for the set Y, which is a point or a segment).

For $G(x)$, $G_1(x)$, $G_2(x) \in P^m$ with the areas of definition X, X_1, X_2 respectively, we introduce the following notations.

- G^{-1} is a point-multiple mapping, the graph of which is a reflection of the mapping G graph relative to the line $g = x$. Obviously, if $G \in P^{\uparrow}$ ($G \in P^{\downarrow}$), then $G^{-1} \in P^{\uparrow}$ ($G^{-1} \in P^{\downarrow}$).
- $-G$ is a point-multiple mapping, the graph of which is a reflection of the mapping G graph relative to the line $g = 0$. Moreover, if $G \in P^{\uparrow}$ ($G \in P^{\downarrow}$), then $-G \in P^{\downarrow}$ ($-G \in P^{\uparrow}$).
- $G_1 + G_2$ is a point-set map with a domain of definition $X_1 \cap X_2$, defined as follows: $(G_1 + G_2)(x) = G_1(x) + G_2(x)$ for any $x \in X_1 \cap X_2$. In the case of $X_1 \cap X_2 \neq \emptyset$ the following property is met: if $G_1 \in P^{\uparrow}$, $G_2 \in P^{\uparrow}$ ($G_1 \in P^{\downarrow}$, $G_2 \in P^{\downarrow}$), then $G_1 + G_2 \in P^{\uparrow}$ ($G_1 + G_2 \in P^{\downarrow}$).
- $G_1 - G_2 \equiv G_1 + (-G_2)$.
- $\int\limits_a^b G(x)dx \equiv \int\limits_a^b \min_m(G(x))dx$ is a generalization of the integral for point-multiple mapping where $\min_m(Y) \equiv \{y \in Y \mid |y| \leq |y'| \ \forall\ y' \in Y\}$ is the minimum modulo number from the set $Y \subseteq \mathbb{R}$.

Let's denote by f' the subdifferential[11] for an arbitrary continuous function $f(x)$ defined on a convex set $X \subseteq \mathbb{R}$. Moreover, if f is a convex (concave) function, then $f' \in P^{\uparrow}$ ($f' \in P^{\downarrow}$).

[11] The subdifferential is a point-set mapping and is used instead of the derivative in cases where the latter may not exist (Vasiliev F. P., 1988, [54]). The existence of one-sided derivatives is sufficient for the existence of a subdifferential.

1.3.2. Supply function

The Walras <u>*supply function*</u> (Vasin A. A., Morozov V. V., 2005, [51])

$$S_i(p_i) \equiv \underset{v_i \in [0, V_i^{\max}]}{\text{Arg max}} \, (p_i v_i - c_i(v_i))$$

is related to the production cost function $c_i(v_i)$ and determines the optimal output volume of node i producers in terms of maximizing their profits at a fixed selling price p_i. This function is a non-decreasing regular point-set mapping (i.e. $S_i(p_i) \in P^{\uparrow}$) defined at $p_i \geq 0$, with $\min(S_i(0)) = 0$. Thus, $S_i(p_i)$ is a point or a segment for any $p_i \geq 0$.

The following relations are derived from the definition of the supply function, linking the functions $c_i(v_i)$ and $S_i(p_i)$:

$$S_i(p_i) = \begin{cases} \{0\}, & V_i^{\max} = 0, \\ \begin{cases} \{0\}, & p_i \in [0, c'_{i+}(0)), \\ (c'_i)^{-1}(p_i), & p_i \in [c'_{i+}(0), c'_{i-}(V_i^{\max})], \\ \{V_i^{\max}\}, & p_i \in (c'_{i-}(V_i^{\max}), +\infty), \end{cases} & V_i^{\max} > 0, \end{cases}$$

$$c_i(v_i) = \int_0^{v_i} S_i^{-1}(u) du.$$

Figure 4 shows an example of a piecewise linear production cost function $c_i(v_i)$ together with its corresponding supply function $S_i(p_i)$. This example describes a case in which node producers have three production facilities available, each of which allows them to produce a limited amount of goods at a fixed marginal cost.

1.3.3. Demand function

The <u>*demand function*</u>

$$D_i(p_i) \equiv \underset{v_i^d \in [0, V_i^{d,\max}]}{\text{Arg max}} \left(U_i(v_i^d) - p_i v_i^d \right)$$

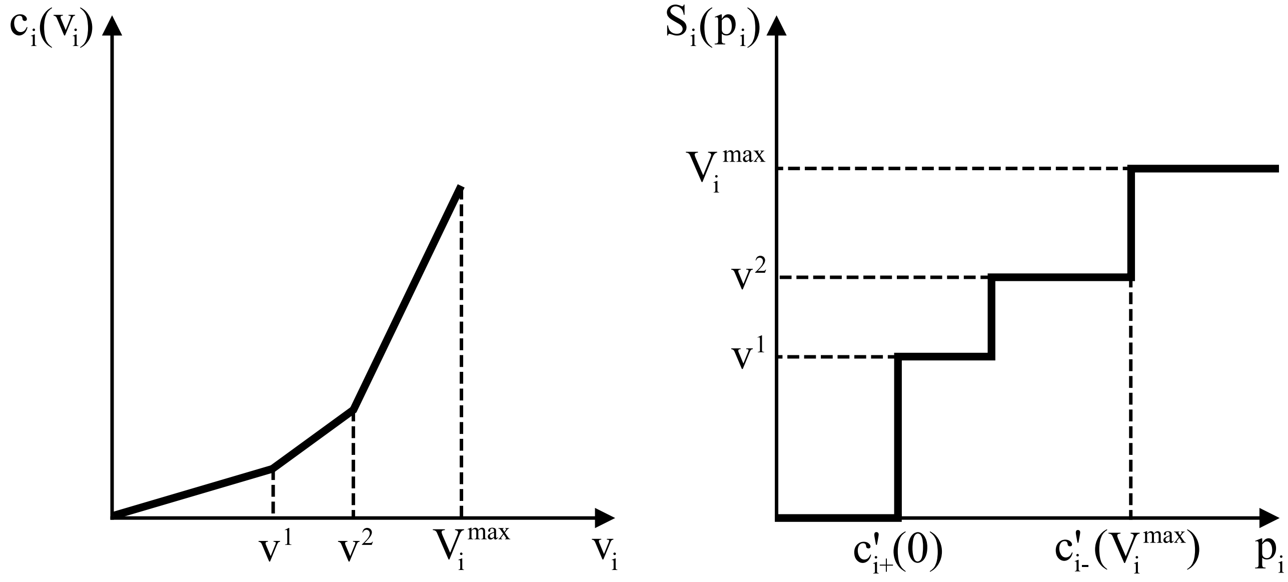

Fig. 4. An example of a piecewise linear production cost function $c_i(v_i)$ (left) and the corresponding supply function $S_i(p_i)$ (right)

is related to the utility function $U_i(v_i^d)$ and determines the optimal purchase volume for consumers of node i in terms of maximizing their gains at a fixed price for the product p_i. This function is a non-increasing regular point-set mapping (i.e. $D_i(p_i) \in P^{\downarrow}$) defined at $p_i \geq 0$ and equal to zero for sufficiently large p_i, with $D_i(0)$ being a point.

The following relations are derived from the definition of the demand function, linking the functions $U_i(v_i^d)$ and $D_i(p_i)$:

$$D_i(p_i) = \begin{cases} \{V_i^{d,\max}\}, & p_i = 0, \\ (U_i')^{-1}(p_i), & p_i \in (0, U_{i+}'(0)], \\ \{0\}, & p_i \in (U_{i+}'(0), +\infty), \end{cases} \qquad U_i(v_i^d) = \int\limits_0^{v_i^d} D_i^{-1}(u)du. \tag{1.11}$$

Figure 5 shows an example of a piecewise linear utility function of consumption $U_i(v_i^d)$ together with its corresponding demand function $D_i(p_i)$. This example corresponds to the case in which, for consumers of the node, the utility of purchasing another unit of goods does not change with an increase in consumption until the saturation volume $V^{d,\max}$ is reached.

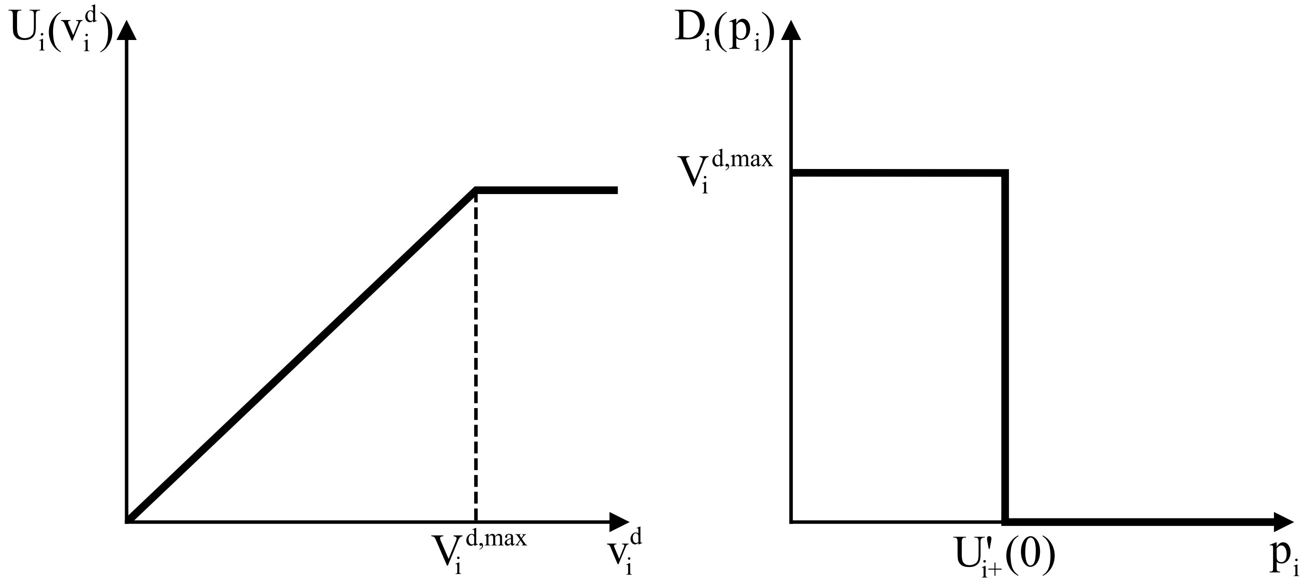

Fig. 5. An example of a piecewise linear utility function of consumption $U_i(v_i^d)$ (left) and its corresponding demand function $D_i(p_i)$ (right)

1.3.4. Functions of marginal transmission costs

Let's introduce for each line $\{i, j\} \in L$ *functions of marginal transmission costs* $e_{ij}^{nex}(q_{ij})$ and $e_{ij}^{ex}(q_{ij})$. The first function corresponds to the case of no line extension, the second to the case of line extension. The function $e_{ij}^{nex}(q_{ij})$ is defined as follows: if $Q^0_{\{i,j\}} = 0$, then $e_{ij}^{nex}(q_{ij})$ is defined only when $q_{ij} = 0$ and is equal to $(-\infty, +\infty)$ at this point; otherwise

$$e_{ij}^{nex}(q_{ij}) = \begin{cases} \big(-\infty, -e^t_{\{i,j\}}\big], & q_{ij} = -Q^0_{\{i,j\}}, \\ \{-e^t_{\{i,j\}}\}, & q_{ij} \in (-Q^0_{\{i,j\}}, 0), \\ \big[-e^t_{\{i,j\}}, e^t_{\{i,j\}}\big], & q_{ij} = 0, \\ \{e^t_{\{i,j\}}\}, & q_{ij} \in (0, Q^0_{\{i,j\}}), \\ \big[e^t_{\{i,j\}}, +\infty\big), & q_{ij} = Q^0_{\{i,j\}}. \end{cases}$$

The function $e^{ex}_{ij}(q_{ij})$ is defined as follows: if $\Delta Q^{\max}_{\{i,j\}} = 0$, then $e^{ex}_{ij}(q_{ij}) \equiv e^{nex}_{ij}(q_{ij})$; if $\Delta Q^{\max}_{\{i,j\}} \neq 0$, $Q^0_{\{i,j\}} = 0$, then

$$e^{ex}_{ij}(q_{ij}) = \begin{cases} \big(-\infty, -e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'_{-}(\Delta Q^{\max}_{\{i,j\}})\big], & q_{ij} = -Q^{\max}_{\{i,j\}}, \\ -e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'(|q_{ij}| - Q^0_{\{i,j\}}), & q_{ij} \in (-Q^{\max}_{\{i,j\}}, 0), \\ \big[-e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'_{+}(0), e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'_{+}(0)\big], & q_{ij} = 0, \\ e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'(|q_{ij}| - Q^0_{\{i,j\}}), & q_{ij} \in (0, Q^{\max}_{\{i,j\}}), \\ \big[e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'_{-}(\Delta Q^{\max}_{\{i,j\}}), +\infty\big), & q_{ij} = Q^{\max}_{\{i,j\}}; \end{cases}$$

if $\Delta Q^{\max}_{\{i,j\}} \neq 0$, $Q^0_{\{i,j\}} \neq 0$, then

$$e^{ex}_{ij}(q_{ij}) = \begin{cases} \big(-\infty, -e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'_{-}(\Delta Q^{\max}_{\{i,j\}})\big], & q_{ij} = -Q^{\max}_{\{i,j\}}, \\ -e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'(|q_{ij}| - Q^0_{\{i,j\}}), & q_{ij} \in (-Q^{\max}_{\{i,j\}}, -Q^0_{\{i,j\}}), \\ \big[-e^t_{\{i,j\}} - {E^v_{\{i,j\}}}'_{+}(0), -e^t_{\{i,j\}}\big], & q_{ij} = -Q^0_{\{i,j\}}, \\ \{-e^t_{\{i,j\}}\}, & q_{ij} \in (-Q^0_{\{i,j\}}, 0), \\ \big[-e^t_{\{i,j\}}, e^t_{\{i,j\}}\big], & q_{ij} = 0, \\ \{e^t_{\{i,j\}}\}, & q_{ij} \in (0, Q^0_{\{i,j\}}), \\ \big[e^t_{\{i,j\}}, e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'_{+}(0)\big], & q_{ij} = Q^0_{\{i,j\}}, \\ e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'(|q_{ij}| - Q^0_{\{i,j\}}), & q_{ij} \in (Q^0_{\{i,j\}}, Q^{\max}_{\{i,j\}}), \\ \big[e^t_{\{i,j\}} + {E^v_{\{i,j\}}}'_{-}(\Delta Q^{\max}_{\{i,j\}}), +\infty\big), & q_{ij} = Q^{\max}_{\{i,j\}}. \end{cases}$$

Each of these functions is a non-decreasing regular point-set mapping (i.e. $e^{nex}_{ij}(q_{ij}), e^{ex}_{ij}(q_{ij}) \in P^{\uparrow}$) with a graph passing through the origin and symmetrical with respect to it (figure 6).

1.3.5. Competitive equilibrium

We introduce an important concept of competitive equilibrium, equivalent to the corresponding definition from [43] (Vasin A. A., Grigorieva O. M., Tsyganov N. I., 2019).

Definition 4. *For a fixed set of expanded lines $R \subseteq L$, the combination of the prices vector $\vec{p} = (p_i \mid i \in N)$, production volumes vector $\vec{v} = (v_i \mid i \in N)$ and flows vector $\vec{q} = (q_{ij} \mid \{i,j\} \in L)$*

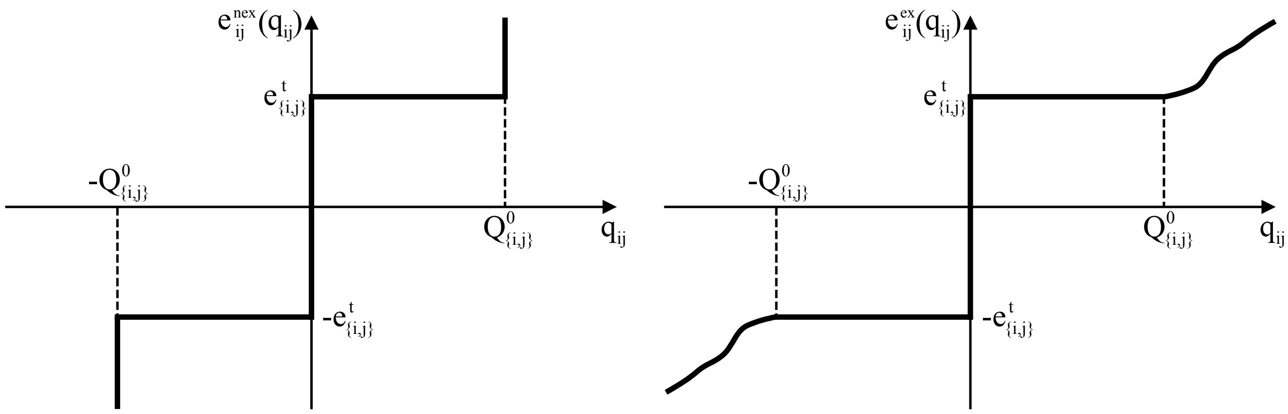

Fig. 6. Example of marginal transmission cost functions $e_{ij}^{nex}(q_{ij})$ (left) and $e_{ij}^{ex}(q_{ij})$ (right) with $\Delta Q_{\{i,j\}}^{\max} = +\infty$

is called the competitive equilibrium of the market if it meets the following conditions:

$$v_i \in S_i(p_i),\ v_i^d(\vec{q},\vec{v}) \in D_i(p_i) \textit{ for any node } i \in N; \tag{1.12}$$

$$p_j - p_i \in \begin{cases} e_{ij}^{nex}(q_{ij}), & \textit{if } \{i,j\} \notin R, \\ e_{ij}^{ex}(q_{ij}), & \textit{if } \{i,j\} \in R \end{cases} \textit{ for any line } \{i,j\} \in L. \tag{1.13}$$

Let's denote the set of all triples $(\vec{p},\vec{v},\vec{q})$ meeting (1.12, 1.13) by $\widetilde{X}^{comp}(R)$, and we will call the corresponding prices $\vec{p}$, production volumes $\vec{v}$, flows $\vec{q}$ and consumption volumes $\vec{v^d}(\vec{q},\vec{v})$ equilibrium.

The first condition means that in each node $i \in N$ at a fixed price p_i the volume of production v_i maximizes producers' profit, the volume of consumption $v_i^d(\vec{q},\vec{v})$ maximizes the gain of consumers, and the price p_i balances supply and demand in the node, taking into account incoming and outgoing flows. The second condition says that for each line $\{i,j\} \in L$ at fixed prices p_i and p_j at incident nodes the flow value $|q_{ij}|$ from a node with a lower price to a node with a higher price reached such a value at which the marginal transport costs equaled the price difference, making further increasing the flow is unprofitable for the owner of the line $\{i,j\}$. In other words, competitive equilibrium is a market condition in which, at fixed prices, it is not profitable for any of the market players to change their behaviour.

The following statement concretizing the well-known welfare theorem (Arrow K. D. and Debre J., 1954, [7]) and proved in [43] (Vasin A. A., Grigorieva O. M., Tsyganov N. I., 2019), establishes a connection between competitive equilibrium and the solution of the auxiliary problem (1.6).

Theorem 2. *Task (1.6) is a convex programming problem, and* $(\vec{q}, \vec{v}) \in \widetilde{X}(R)$ *if and only if* $\exists\ \vec{p}$: $(\vec{p}, \vec{v}, \vec{q}) \in \widetilde{X}^{comp}(R)$.

Thus, to solve the auxiliary problem (1.6), which is now reduced to the solution (1.12, 1.13), standard convex optimization methods can be used (Sukharev A. G., Timokhov A. V., Fedorov V. B., 1985, [55]). In the next paragraph, a special algorithm for solving this problem is proposed.

The following important conclusion also follows from theorem 2. Let L^* be the optimal set of extensible lines found according to (1.7), and $(\overrightarrow{p^*}, \overrightarrow{v^*}, \overrightarrow{q^*}) \in \widetilde{X}^{comp}(L^*)$. Then, when encouraging line owners to make decisions to «expand/not expand» in accordance with L^* and the inability of players to influence prices, the state $(\overrightarrow{p^*}, \overrightarrow{v^*}, \overrightarrow{q^*})$ that ensures maximum public welfare is stable, i.e. rational players have no incentive to deviate from it. Line owners can be stimulated, for example, by compensating fixed costs $E^f_{\{i,j\}}$ on the part of a player interested in implementing an optimal condition. As previously noted, this player may be the state.

1.3.6. Algorithm for solving the auxiliary problem

Let's describe the algorithm for solving the auxiliary problem (1.6). Its main idea is to transfer the aggregated net supply from all nodes to some node $i_0 \in N$, calculate the equilibrium price for this node and then reverse the movement, in the course of which a competitive equilibrium $(\overrightarrow{\tilde{p}}, \overrightarrow{\tilde{v}}, \overrightarrow{\tilde{q}})$ meeting (1.12, 1.13) is calculated. According to theorem 2, the found vectors of flows $\overrightarrow{\tilde{q}}$ and production volumes $\overrightarrow{\tilde{v}}$ form the solution of the problem (1.6).

The **algorithm** consists of the following three steps.

Step 1. Building a root tree of minimum height. Let's choose node $i_0 \in N$ in such a way that the root tree corresponding to graph $G = (N, L)$ with root i_0 has the minimum possible height h. Let's denote by $\sigma(i)$ the mapping that matches an arbitrary node $i \in N$ with its nearest predecessor (parent), while we assume that $\sigma(i_0) = i_0$. We divide the set of all nodes N into subsets N_1, ..., N_{h+1} as follows:

- $N_1 = \{i \in N \mid \sigma^{-1}(i) = \emptyset\}$ - the set of final nodes;
- $N_k = \left\{ i \in N \setminus \bigcup_{l=1}^{k-1} N_l \;\middle|\; \sigma^{-1}(i) \subseteq \bigcup_{l=1}^{k-1} N_l \right\}$, $k = 2$, ..., $h+1$.

In this case, $N_{h+1} = \{i_0\}$. Figures 7 and 8 show an example of the transport structure of the market and the corresponding root tree of minimum height with $h = 3$ and $i_0 = 6$. Let's describe one of the ways to calculate the values h, i_0, $\sigma(i)$, N_k for given sets of adjacent nodes $(Z(i) \mid i \in N)$, whose complexity is $O(|N|^2)$. The desired tree is built by sequentially including nodes, starting with the final ones. Let's introduce the following values: $\widehat{N}$ - the set of nodes not yet included; $\widehat{Z}_i$ - the set of nodes not yet included adjacent to the node i; $\widehat{d}_i$ - the number of nodes not yet included adjacent to node i. For an arbitrary singleton set A let's denote by $J(A)$ its only element.

Substep 1.0. Suppose $\widehat{N} = N$, $\widehat{Z}_i = Z(i)$, $\widehat{d}_i = |Z(i)|$ for each $i \in N$.

Substep 1.k, k=1, ... If $|\widehat{N}| > 2$, then we assume:

1) $N_k = \{i \in \widehat{N} \mid \widehat{d}_i = 1\}$, $\widehat{N} = \widehat{N} \setminus N_k$;
2) for every $i \in N_k$: $\sigma(i) = J(\widehat{Z}_i)$, $\widehat{Z}_{J(\widehat{Z}_i)} = \widehat{Z}_{J(\widehat{Z}_i)} \setminus \{i\}$, $\widehat{d}_{J(\widehat{Z}_i)} = \widehat{d}_{J(\widehat{Z}_i)} - 1$.

If $|\widehat{N}| \leq 2$, then we assume:

$$\begin{cases} h = k-1, i_0 = i_1, N_{h+1} = \{i_0\}, \sigma(i_0) = i_0, & \text{if } \widehat{N} = \{i_1\}, \\ N_k = \{i_1\}, \sigma(i_1) = i_2, h = k, i_0 = i_2, N_{h+1} = \{i_0\}, \sigma(i_0) = i_0, & \text{if } \widehat{N} = \{i_1, i_2\}, \end{cases}$$

after that, step 1 is completed.

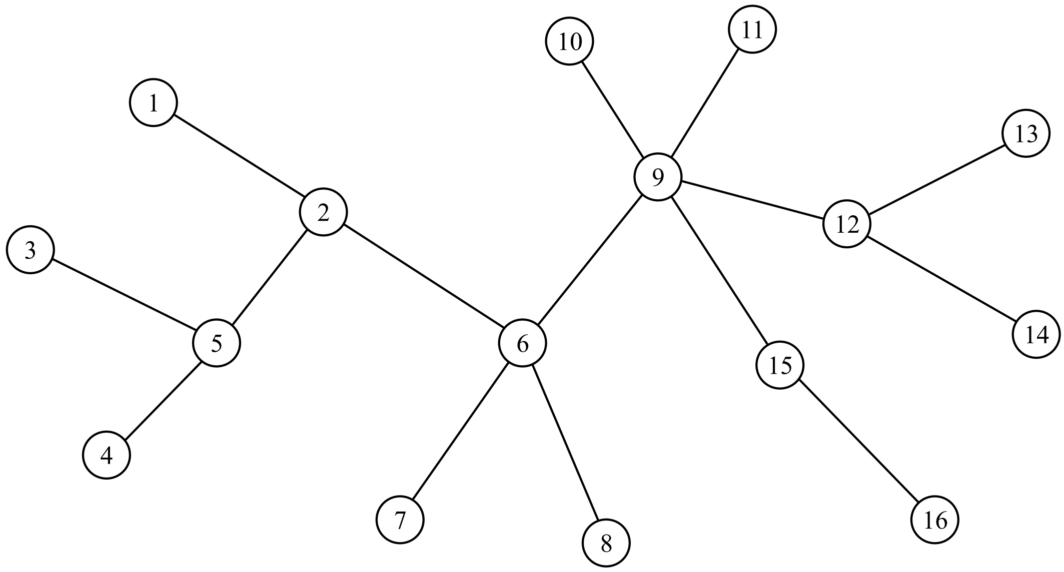

Fig. 7. An example of a transport market structure

Step 2. Transferring the balance of supply and demand to the root node (direct move). Denote by $\Delta S_i(p_i) = S_i(p_i) - D_i(p_i)$ the function of pure supply in the node $i \in N$, showing the excess of supply over demand at a fixed price in the node, without taking into account the transfer of goods between nodes.

We introduce the following two functions: $\Delta\overline{S}_j(p_j)$ and $\Delta S_{ij}(p_j)$. The function $\Delta\overline{S}_j(p_j)$ determines the aggregated net supply at node $j \in N$, taking into account the transfer of goods between this node and subsequent nodes, i.e. the excess of supply over demand for a subtree with root j at a fixed price at this node. The function $\Delta S_{ij}(p_j)$ determines the flow from node $i \in N$ to the preceding node $j = \sigma(i)$, which, at a fixed price in node j balances supply and demand in

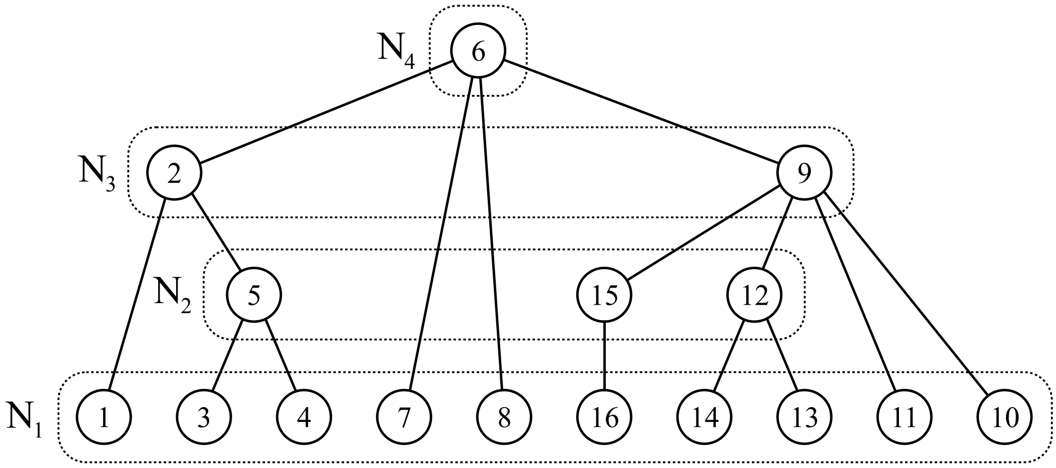

Fig. 8. Root tree of the minimum height for the market shown in figure 7

the subtree with the root vertex i and at the same time ensures maximum profit for the owner of the line $\{i, j\}$ provided that it cannot influence the price at node i.

At this step, the functions $\Delta\overline{S}_j(p_j)$ and $\Delta S_{j\sigma(j)}(p_{\sigma(j)})$ are sequentially calculated, starting from the final nodes and ending with the root node.

Substep 2.k, k=1, ..., h. For each node $j \in N_k$, the functions $\Delta\overline{S}_j(p_j)$, $\Delta S_{j\sigma(j)}(p_{\sigma(j)})$ are calculated according to the operations introduced in paragraph 1.3.1.:

$$\Delta\overline{S}_j(p_j) = \Delta S_j(p_j) + \sum_{i\in\sigma^{-1}(j)} \Delta S_{ij}(p_j); \tag{1.14}$$

$$\Delta S_{j\sigma(j)}(p_{\sigma(j)}) = \begin{cases} ((\Delta\overline{S}_j)^{-1} + e^{nex}_{j\sigma(j)})^{-1}(p_{\sigma(j)}), & \{j, \sigma(j)\} \notin R, \\ ((\Delta\overline{S}_j)^{-1} + e^{ex}_{j\sigma(j)})^{-1}(p_{\sigma(j)}), & \{j, \sigma(j)\} \in R. \end{cases} \tag{1.15}$$

Substep 2.(h+1). The function $\Delta\overline{S}_{i_0}(p_{i_0})$ is calculated for the root node:

$$\Delta\overline{S}_{i_0}(p_{i_0}) = \Delta S_{i_0}(p_{i_0}) + \sum_{i\in\sigma^{-1}(i_0)} \Delta S_{ii_0}(p_{i_0}). \tag{1.16}$$

Step 3. Calculation of competitive equilibrium (reverse). At this step, equilibrium prices $\widetilde{p}_i$, production volumes $\widetilde{v}_i$, consumption volumes $\widetilde{v}^d_i$ and flows $\widetilde{q}_{ij}$ are sequentially calculated, starting from the root node and ending with the final ones. Auxiliary coefficients λ_j are also calculated, which are used for ambiguous determination of the basic quantities using the functions $S_j(p_j)$, $D_j(p_j)$, $\Delta S_{ij}(p_j)$ and ensure the fulfillment of condition (1.12) in this case.

Substep 3.1. The following values are calculated for the root node according to the operations entered in paragraph 1.3.1.:

1) $\widetilde{p}_{i_0} = \min((\Delta\overline{S}_{i_0})^{-1}(0))$ - equilibrium price at the root node;

2) $\lambda_{i_0} = \begin{cases} 0, & q^{\min} = q^{\max}, \\ \dfrac{-q^{\min}}{q^{\max} - q^{\min}}, & q^{\min} < q^{\max} \end{cases}$ - auxiliary coefficient for root node where $q^{\min} = \min(\Delta\overline{S}_{i_0}(\widetilde{p}_{i_0}))$, $q^{\max} = \max(\Delta\overline{S}_{i_0}(\widetilde{p}_{i_0}))$;

3) $\widetilde{v}_{i_0} = (1 - \lambda_{i_0}) \cdot \min(S_{i_0}(\widetilde{p}_{i_0})) + \lambda_{i_0} \cdot \max(S_{i_0}(\widetilde{p}_{i_0}))$ - equilibrium volume of production at the root node;

4) $\widetilde{v}^d_{i_0} = \lambda_{i_0} \cdot \min(D_{i_0}(\widetilde{p}_{i_0})) + (1 - \lambda_{i_0}) \cdot \max(D_{i_0}(\widetilde{p}_{i_0}))$ - equilibrium volume of consumption at the root node.

Figure 9 shows an example of calculating and using the auxiliary coefficient λ_{i_0} in substep 3.1.

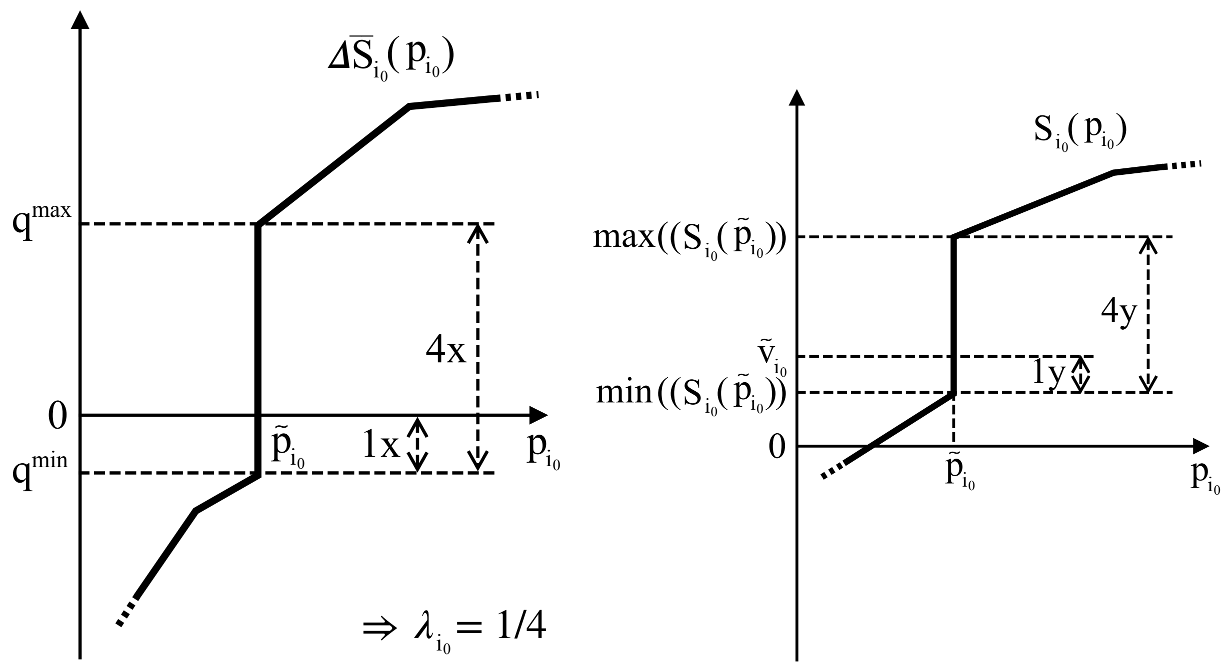

Fig. 9. An example of calculating the auxiliary coefficient λ_{i_0} in item 2 of substep 3.1 (left) and using it to calculate the equilibrium volume of production $\widetilde{v}_{i_0}$ in item 3 of the same substep (on the right); for the functions shown in the graphs $\lambda_{i_0} = \frac{-q^{\min}}{q^{\max} - q^{\min}} = 1/4$, therefore $\widetilde{v}_{i_0} = 3/4 \cdot \min(S_{i_0}(\widetilde{p}_{i_0})) + 1/4 \cdot \max(S_{i_0}(\widetilde{p}_{i_0}))$;

Substep 3.m, m=2, ..., h+1. The set N_k is considered, where $k = h - m + 2$. The following values are calculated for each node $i \in N_k$:

1) $\widetilde{q}_{i\sigma(i)} = (1 - \lambda_{\sigma(i)}) \cdot \min(\Delta S_{i\sigma(i)}(\widetilde{p}_{\sigma(i)})) + \lambda_{\sigma(i)} \cdot \max(\Delta S_{i\sigma(i)}(\widetilde{p}_{\sigma(i)}))$ - the equilibrium flow from node i to node $\sigma(i)$;

2) $\widetilde{p}_i = \begin{cases} \min\left((\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) \cap \left(p_{\sigma(i)} - e^{nex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)})\right)\right), & \{i, \sigma(i)\} \notin R, \\ \min\left((\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) \cap \left(p_{\sigma(i)} - e^{ex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)})\right)\right), & \{i, \sigma(i)\} \in R \end{cases}$ - equilibrium price in node i;

3) $\lambda_i = \begin{cases} 0, & q^{\min} = q^{\max}, \\ \dfrac{\widetilde{q}_{i\sigma(i)} - q^{\min}}{q^{\max} - q^{\min}}, & q^{\min} < q^{\max} \end{cases}$ - the auxiliary coefficient for node i, where $q^{\min} = \min(\Delta\overline{S}_i(\widetilde{p}_i))$, $q^{\max} = \max(\Delta\overline{S}_i(\widetilde{p}_i))$;

4) $\widetilde{v}_i = (1-\lambda_i)\cdot\min(S_i(\widetilde{p}_i)) + \lambda_i\cdot\max(S_i(\widetilde{p}_i))$ - the equilibrium volume of production at node i;

5) $\widetilde{v}_i^d = \lambda_i\cdot\min(D_i(\widetilde{p}_i)) + (1-\lambda_i)\cdot\max(D_i(\widetilde{p}_i))$ - the equilibrium volume of consumption in node i.

The main amount of calculations falls on the second step of the algorithm, where operations on functions are performed. The introduction of the minimum height requirement for the root tree used in the algorithm is related to the possibility of independent calculations of functions $\Delta\overline{S}_j(p_j)$, $\Delta S_{j\sigma(j)}(p_{\sigma(j)})$ in the second step for nodes j belonging to the same set N_k, which, when using parallel computing systems, reduces the time to solve the problem. For the third step, similar reasoning is valid.

Theorem 3. *All calculations of the algorithm are correct*[12]*. The vectors of prices $\overrightarrow{\widetilde{p}} = (\widetilde{p}_i \mid i \in N)$, production volumes $\overrightarrow{\widetilde{v}} = (\widetilde{v}_i \mid i \in N)$ and flows $\overrightarrow{\widetilde{q}} = (\widetilde{q}_{ij} \mid \{i,j\} \in L)$, resulting from the work of the algorithm, meet the following relations: $(\overrightarrow{\widetilde{p}}, \overrightarrow{\widetilde{v}}, \overrightarrow{\widetilde{q}}) \in \widetilde{X}^{comp}(R)$, $(\overrightarrow{\widetilde{q}}, \overrightarrow{\widetilde{v}}) \in \widetilde{X}(R)$.*

Thus, the algorithm finds a competitive equilibrium (1.12, 1.13) and a solution to the auxiliary problem (1.6).

Proof. *Since $S_i \in P^{\uparrow}$, $D_i \in P^{\downarrow}$ for any node $i \in N$, then $\Delta S_i \in P^{\uparrow}$, and $\Delta S_i(p_i)$ is defined if $p_i \geq 0$, is non-negative for sufficiently large p_i, $\Delta S_i(p_i)$ is a point or segment for any $p_i \geq 0$ and $\min(\Delta S_i(0)) \leq 0$. It is proved by induction that all functions of the form $\Delta\overline{S}_i(p_i)$ and $\Delta S_{i\sigma(i)}(p_{\sigma(i)})$ calculated at the second step of the algorithm according to (1.14-1.16) also satisfy similar properties. This takes into account the properties of the operations and functions $e^{nex}_{i\sigma(i)}(q_{i\sigma(i)})$, $e^{ex}_{i\sigma(i)}(q_{i\sigma(i)})$ introduced in paragraph 1.3.1..*

For convenience, we introduce a dummy variable $\widetilde{q}_{i_0 i_0} = 0$ for the root node. We prove by induction that all calculations of the third step of the algorithm are correct, and the obtained values satisfy the following relations:

$$\widetilde{v}_i \in S_i(\widetilde{p}_i),\ \widetilde{v}_i^d \in D_i(\widetilde{p}_i),\ \widetilde{p}_i \geq 0,\ \lambda_i \in [0,1],\ \widetilde{q}_{i\sigma(i)} \in \Delta\overline{S}_i(\widetilde{p}_i) \text{ for any } i \in N; \tag{1.17}$$

[12] The correctness of calculations is understood as the absence of the following operations during the work of the algorithm: dividing by zero, finding the minimum or maximum of an empty set.

$$\widetilde{p}_{\sigma(i)} - \widetilde{p}_i \in \begin{cases} e^{nex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)}), & \textit{if } \{i,\sigma(i)\} \notin R, \\ e^{ex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)}), & \textit{if } \{i,\sigma(i)\} \in R \end{cases} \quad \textit{for any } i \in N \setminus \{i_0\}. \tag{1.18}$$

Since for any $i \in N$, $p_i \geq 0$ each of the value $\Delta\overline{S}_i(p_i)$, $\Delta S_{i\sigma(i)}(p_{\sigma(i)})$, $S_i(p_i)$, $D_i(p_i)$ is defined and is a point or a segment, then the min *and* max *operations applied in the algorithm for such expressions are correct.*

Let's consider substep 3.1. Since $\Delta\overline{S}_{i_0} \in P^{\uparrow}$, $\min(\Delta\overline{S}_{i_0}(0)) \leq 0$ and $\Delta\overline{S}_{i_0}(p_{i_0})$ is non-negative for sufficiently large p_{i_0}, then $(\Delta\overline{S}_{i_0})^{-1}(0) \cap [0,+\infty) \neq \emptyset$, therefore, the action 1) of this substep is correct, and $\widetilde{p}_{i_0} \geq 0$ and $0 \in \Delta\overline{S}_{i_0}(\widetilde{p}_{i_0})$. It follows from the last property that $\min(\Delta\overline{S}_{i_0}(\widetilde{p}_{i_0})) \leq 0$, $\max(\Delta\overline{S}_{i_0}(\widetilde{p}_{i_0})) \geq 0$, which means $\lambda_{i_0} \in [0,1]$, which, in turn, means the validity of inclusions $\widetilde{v}_{i_0} \in S_{i_0}(\widetilde{p}_{i_0})$ and $\widetilde{v}^d_{i_0} \in D_{i_0}(\widetilde{p}_{i_0})$.

Let the relations (1.17, 1.18) be true for the values calculated in substeps 3.1, ..., 3.(m-1). Let's consider substep 3.m. Let's fix an arbitrary node $i \in N_{h-m+2}$. By the assumption of induction $\widetilde{p}_{\sigma(i)} \geq 0$, $\lambda_{\sigma(i)} \in [0,1]$, therefore $\widetilde{q}_{i\sigma(i)} \in \Delta S_{i\sigma(i)}(\widetilde{p}_{\sigma(i)})$, which in the case of $\{i,\sigma(i)\} \notin R$ is equivalent to

$$\widetilde{q}_{i\sigma(i)} \in \left((\Delta\overline{S}_i)^{-1} + e^{nex}_{i\sigma(i)}\right)^{-1}(\widetilde{p}_{\sigma(i)}) \Leftrightarrow \widetilde{p}_{\sigma(i)} \in \left((\Delta\overline{S}_i)^{-1} + e^{nex}_{i\sigma(i)}\right)(\widetilde{q}_{i\sigma(i)}) \Leftrightarrow$$
$$\Leftrightarrow \widetilde{p}_{\sigma(i)} \in (\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) + e^{nex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)}) \Leftrightarrow 0 \in (\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) - \left(\widetilde{p}_{\sigma(i)} - e^{nex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)})\right) \Leftrightarrow$$
$$\Leftrightarrow (\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) \cap \left(\widetilde{p}_{\sigma(i)} - e^{nex}_{i\sigma(i)}(\widetilde{q}_{i\sigma(i)})\right) \neq \emptyset.$$

Note that the function $\Delta\overline{S}_i(p_i)$ is not defined for $p_i < 0$, therefore $(\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)}) \subseteq [0,+\infty)$, this means that for $\{i,\sigma(i)\} \notin R$ action 2) of this substep is correct, and $\widetilde{p}_i \geq 0$. In the case of $\{i,\sigma(i)\} \in R$, the situation is similar. Also, this action implies the fulfillment of condition (1.18) for node i and the inclusion of $\widetilde{p}_i \in (\Delta\overline{S}_i)^{-1}(\widetilde{q}_{i\sigma(i)})$. From the latter it follows that $\widetilde{q}_{i\sigma(i)} \in \Delta\overline{S}_i(\widetilde{p}_i) \Leftrightarrow \widetilde{p}_i \in \left[\min(\Delta\overline{S}_i(\widetilde{p}_i)), \max(\Delta\overline{S}_i(\widetilde{p}_i))\right]$, therefore $\lambda_i \in [0,1]$, which means the validity of the ratios $\widetilde{v}_i \in S_i(\widetilde{p}_i)$ and $\widetilde{v}^d_i \in D_i(\widetilde{p}_i)$. Thus, the induction assumption is proved for step 3.m and properties (1.17, 1.18) are met.

Now let's prove that

$$\widetilde{v}^d_i = v^d_i(\overrightarrow{\widetilde{q}}, \overrightarrow{\widetilde{v}}) \textit{ for any } i \in N. \tag{1.19}$$

Consider the difference $v^d_i(\overrightarrow{\widetilde{q}}, \overrightarrow{\widetilde{v}}) - \widetilde{v}^d_i$ for an arbitrary node $i \in N$ and show that it is equal to *zero. Indeed, denoting* $q^{\min} = \min(\Delta\overline{S}_i(\widetilde{p}_i))$, $q^{\max} = \max(\Delta\overline{S}_i(\widetilde{p}_i))$, *we get:*

$$v_i^d(\overrightarrow{\widetilde{q}}, \overrightarrow{\widetilde{v}}) - \widetilde{v}_i^d = \widetilde{v}_i - \sum_{j\in Z(i)} \widetilde{q}_{ij} - \widetilde{v}_i^d = \widetilde{v}_i - \widetilde{v}_i^d + \sum_{j\in\sigma^{-1}(i)} \widetilde{q}_{ji} - \widetilde{q}_{i\sigma(i)} =$$

$$= [(1-\lambda_i)\cdot\min(S_i(\widetilde{p}_i)) + \lambda_i\cdot\max(S_i(\widetilde{p}_i))] - [\lambda_i\cdot\min(D_i(\widetilde{p}_i)) + (1-\lambda_i)\cdot\max(D_i(\widetilde{p}_i))] +$$

$$+ \sum_{j\in\sigma^{-1}(i)} [(1-\lambda_i)\cdot\min(\Delta S_{ji}(\widetilde{p}_i)) + \lambda_i\cdot\max(\Delta S_{ji}(\widetilde{p}_i))] - \widetilde{q}_{i\sigma(i)} =$$

$$= (1-\lambda_i)\cdot\left[\min(S_i(\widetilde{p}_i)) + \min(-D_i(\widetilde{p}_i)) + \sum_{j\in\sigma^{-1}(i)} \min(\Delta S_{ji}(\widetilde{p}_i))\right] +$$

$$+ \lambda_i\cdot\left[\max(S_i(\widetilde{p}_i)) + \max(-D_i(\widetilde{p}_i)) + \sum_{j\in\sigma^{-1}(i)} \max(\Delta S_{ji}(\widetilde{p}_i))\right] - \widetilde{q}_{i\sigma(i)} =$$

$$= (1-\lambda_i)\cdot\min(\Delta\overline{S}_i(\widetilde{p}_i)) + \lambda_i\cdot\max(\Delta\overline{S}_i(\widetilde{p}_i)) - \widetilde{q}_{i\sigma(i)} = (1-\lambda_i)\cdot q^{\min} + \lambda_i\cdot q^{\max} - \widetilde{q}_{i\sigma(i)} =$$

$$= \lambda_i\cdot(q^{\max} - q^{\min}) + q^{\min} - \widetilde{q}_{i\sigma(i)} = \begin{cases} q^{\min} - \widetilde{q}_{i\sigma(i)}, & q^{\min} = q^{\max}, \\ (\widetilde{q}_{i\sigma(i)} - q^{\min}) + q^{\min} - \widetilde{q}_{i\sigma(i)}, & q^{\min} < q^{\max} \end{cases} = 0.$$

In these transitions, the ratios (1.3, 1.14, 1.16) and obvious equalities $\widetilde{q}_{ij} = -\widetilde{q}_{ji}$, $-\max(D_i(\widetilde{p}_i)) = \min(-D_i(\widetilde{p}_i))$, $-\min(D_i(\widetilde{p}_i)) = \max(-D_i(\widetilde{p}_i))$ *were used, as well as the property* $\widetilde{q}_{i\sigma(i)} \in \Delta\overline{S}_i(\widetilde{p}_i)$.

Thus, the conditions (1.17-1.19) are met. Note that in this case the obtained values also satisfy the relations (1.12, 1.13), therefore $(\overrightarrow{\widetilde{p}}, \overrightarrow{\widetilde{v}}, \overrightarrow{\widetilde{q}}) \in \widetilde{X}^{comp}(R)$. *Taking into account theorem 2, the inclusion* $(\overrightarrow{\widetilde{q}}, \overrightarrow{\widetilde{v}}) \in \widetilde{X}(R)$ *is also valid.* ■

1.3.7. Estimation of the complexity of the algorithm for the case of piecewise linear functions

Let's call a monotone regular point-multiple mapping $G(x)$ with a domain of definition X piecewise linear if its graph $Gr(G) = \{(x,g) \mid x \in X,\ g \in G(x)\}$ is represented as a union of a finite number of segments and rays. Let's call a point $A \in Gr(G)$ the vertex of a piecewise linear map $G(x)$ if there is no segment $[A_1, A_2] \subseteq Gr(G)$ for which A is an internal point.

Consider a special case of a problem with piecewise linear functions $S_i(p_i)$, $D_i(p_i)$, $e_{ij}^{nex}(q_{ij})$, $e_{ij}^{ex}(q_{ij})$ (or, equivalently, with piecewise quadratic functions $c_i(v_i)$, $U_i(v_i^d)$, $E_{\{i,j\}}^v(\Delta Q_{\{i,j\}})$) and estimate the complexity of the described algorithm for solving the auxiliary problem for this case. It is of particular interest because any of the initial functions $S_i(p_i)$, $D_i(p_i)$, $e_{ij}^{nex}(q_{ij})$, $e_{ij}^{ex}(q_{ij})$ can be approximated with the necessary degree of accuracy by some piecewise linear function, which enables approximate solution of the auxiliary problem (1.6) for arbitrary source functions.

Theorem 4. *Let each of the initial functions $S_i(p_i)$, $D_i(p_i)$, $e_{ij}^{nex}(q_{ij})$, $e_{ij}^{ex}(q_{ij})$ be piecewise linear for any $i \in N$, $\{i,j\} \in L$, and the number of vertices of each such function does not exceed some predefined parameter. Then the complexity of the developed algorithm for solving the auxiliary problem (1.6) is $O(|N|^2)$.*

Proof. *Let's prove that the complexity of step 1 is $O(|N|^2)$. The complexity of operation $\widetilde{N} = \widetilde{N} \setminus N_k$ does not exceed $O(|N_k| \cdot |N|)$, so the total complexity of all such operations does not exceed $O\left(\sum\limits_{k=1}^{h+1} |N_k| \cdot |N|\right) = O(|N|^2)$. It remains to be noted that any other operation of step 1 has a complexity not exceeding $O(|N|)$ and is applied no more than $|N|$ times.*

Let's estimate the complexity of the second step. It is proved by induction that all functions of the form $\Delta S_i(p_i)$, $\Delta\overline{S}_i(p_i)$ and $\Delta S_{i\sigma(i)}(p_{\sigma(i)})$ are also piecewise linear, like the original functions. This follows from the fact that the operations of inversion, addition and subtraction introduced in paragraph 1.3.1., do not derive from the class of piecewise linear functions. Any non-decreasing piecewise linear function $G(x)$ can be described by the trinity $(A^l, (A_1, ..., A_k), A^r)$, where $(A_1, ..., A_k)$ is an ascending set of points $A_l \in Gr(G)$ (for any $A_l = (x_l, g_l)$, $A_{l+1} = (x_{l+1}, g_{l+1})$ the following is met: $x_l \le x_{l+1}$, $g_l \le g_{l+1}$) defining the segments $[A_1, A_2] \subseteq Gr(G)$, ..., $[A_{k-1}, A_k] \subseteq Gr(G)$, and A^l, A^r are the points defining the directions of the left $[A_1A^l) \subseteq Gr(G)$ and the right $[A_kA^r) \subseteq Gr(G)$ rays. If $A^l = A_1$, then we assume that $[A_1A^l) = \emptyset$, i.e. there is no left ray. We assume the same for the right ray. These segments and rays form graph of the function $G(x)$: $Gr(G) = [A_1A^l) \cup \bigcup\limits_{l=1}^{k-1} [A_l, A_{l+1}] \cup [A_kA^r)$. Thus, the non-decreasing piecewise linear function with k vertices can be described by $k+2$ points. Let's denote by M the parameter from the condition of the theorem, which limits the number of vertices from above for any of the original functions. Therefore, any function of the form $S_i(p_i)$, $-D_i(p_i)$, $e_{ij}^{nex}(q_{ij})$, $e_{ij}^{ex}(q_{ij})$ is definitely determined by $M+2$ points, which, we assume, are fed to the algorithm.

Let's show how for arbitrary non-decreasing piecewise linear functions $G(x)$, $G_1(x)$, $G_2(x)$, given by sets of points $(A^l, (A_1, ..., A_k), A^r)$, $(A_1^l, (A_{11}, ..., A_{1k_1}), A_1^r)$, $(A_2^l, (A_{21}, ..., A_{2k_2}), A_2^r)$ accordingly, the functions G^{-1} and G_1+G_2, are calculated, and we estimate the complexity of these operations. To find the function G^{-1}, replace each point $A = (x, g)$ from set $(A^l, (A_1, ..., A_k), A^r)$ with point (g, x), while the complexity of such an operation is $O(k)$, and the number of points in the set of function G^{-1} coincides with the number of points in the set of function G.

To calculate the function $G_1 + G_2$ it is sufficient to make one traversal of sets $(A_1^l, (A_{11}, ..., A_{1k_1}), A_1^r)$ and $(A_2^l, (A_{21}, ..., A_{2k_2}), A_2^r)$ from left to right, simultaneously processing *the points of both sets in ascending order of the abscissa x, thereby defining the vertices of the function G_1+G_2 and forming its set. When processing some point $A=(x,g)$ of a set for one of the original functions, it is checked whether x belongs to the definition domain of the second original function, after which, if this condition is met, the values $m_1=\min(G_1(x))$, $m_2=\min(G_2(x))$, $M_1=\max(G_1(x))$, $M_2=\max(G_2(x))$ are found (because to find these values, it is enough to use the last processed point of each set and the point closest to it $A'=(x',g')$, for which $x'>x$, then the complexity of such an operation is $O(1)$). Then $(G_1+G_2)(x)=[m_1+m_2, M_1+M_2]$, so two points are added to the desired set (x, m_1+m_2) and (x, M_1+M_2), if $m_1+m_2<M_1+M_2$, or one point (x, m_1+m_2), if $m_1+m_2=M_1+M_2$. Also, when forming the desired set, it is necessary to correctly determine the extreme points that set the directions of the rays, which also happens during a single crawl. Thus, the complexity of the summation operation is $O(k_1+k_2)$, and the number of points in the desired set, it can be noted, does not exceed k_1+k_2+6 or $k_1'+k_2'+2$, where $k_l'=k_l+2$ is the number of points in the set of the original function $G_l(x)$, $l\in\{1,2\}$. Next, we assume that the functions in the second step of the algorithm are calculated in the described way.*

According to the above made conclusions, the number of points in the set of any function $\Delta S_i(p_i)=S_i(p_i)+(-D_i(p_i))$ does not exceed $2M+6$. Let's denote by $\Lambda(j)$ the set of nodes included in the subtree with the root $j\in N$: $\Lambda(j)=\{i\in N \mid \exists\, l\geq 0\ :\ \sigma^l(i)=j\}$, and through $\lambda(j)=|\Lambda(j)|$ the number of such nodes. Let $\overline{z}_j$, z_{ij} be the number of points in the sets of functions $\Delta\overline{S}_j(p_j)$ and $\Delta S_{ij}(p_j)$ respectively. We prove by induction that

$$\overline{z}_j \leq \lambda(j)(3M+14)-(M+6), \quad z_{ij}\leq \lambda(i)(3M+14)-2. \tag{1.20}$$

For an arbitrary final node $j\in N_1$, the induction assumption is met, because in this case $z_j\leq 2M+6<2M+8=\lambda(j)(3M+14)-(M+6)$. Let (1.20) be true for $j\in N_1\cup\ldots\cup N_{k-1}$. Consider an arbitrary node $j\in N_k$. Taking into account (1.14, 1.15, 1.20) and the fact that when the function is reversed, the number of points in the set does not change, the following estimates are true:

$$z_{ij}\leq \overline{z}_i+(M+2)+2=\lambda(i)(3M+14)-(M+6)+(M+4)=\lambda(i)(3M+14)-2,\ \ i\in\sigma^{-1}(j);$$

$$\begin{aligned}\overline{z}_j &\leq (2M+6)+\sum_{i\in\sigma^{-1}(j)}(z_{ij}+2)\leq (2M+6)+\sum_{i\in\sigma^{-1}(j)}\lambda(i)(3M+14)=\\ &=(2M+6)+(3M+14)\sum_{i\in\sigma^{-1}(j)}\lambda(i)=(2M+6)+(3M+14)(\lambda(j)-1)=\\ &=\lambda(j)(3M+14)-(M+8)<\lambda(j)(3M+14)-(M+6).\end{aligned}$$

Thus, the assumption of induction is proved for node j, which means that inequalities (1.20) are fulfilled, therefore $\overline{z}_j = O(\lambda(j) \cdot M)$, $z_{ij} = O(\lambda(i) \cdot M)$.

The complexity of the second step of the algorithm is determined by the number of inversion and summation operations in (1.14-1.16), as well as their complexity. The number of inversion operations is $2|N| - 2$*, addition operations is* $3|N| - 2$ *(taking into account the operations used in calculating the functions* $\Delta S_j(p_j)$*). It follows from the proved estimates (1.20) that for any of these operations, the number of points in the set of the resulting function does not exceed* $|N|(3M + 14) - (M + 6)$*, which means that the complexity of the second step does not exceed* $(5|N| - 4) \cdot O(|N|(3M + 14) - (M + 6)) = O(|N|^2)$.

The complexity of each of the $O(|N|)$ *operations of the third step does not exceed* $O(\log_2(|N|))$*, so the complexity of this step is* $o(|N|^2)$. ■

The proved quadratic estimation of the complexity of the developed algorithm for solving the auxiliary problem (1.6) enables asserting its advantage over standard iterative methods[13] of convex optimization.

1.4. Complementary and competitive lines. The flow structure invariance condition

Let's return to the original problem (1.5), which reduces to the problem of finding the optimal set of expanded lines (1.7). Let's introduce some additional concepts that will be needed when building an algorithm for solving this problem.

Definition 5. *Line l is called complementary (respectively competitive) to line $r \neq l$, if $\forall\ M \subseteq L \setminus \{l, r\}$ the inequality*

$$\widetilde{W}((M \cup \{l\}) \cup \{r\}) - \widetilde{W}(M \cup \{l\}) \ \geq (\leq)\ \widetilde{W}(M \cup \{r\}) - \widetilde{W}(M)$$

holds. It follows from the definition that the relations of complementarity and competitiveness are symmetric, i.e. if the line l is complementary (competitive) to line r, then r is complementary (competitive) to l. Let $L_+(l)$ and $L_-(l)$ be sets of complementary and competitive lines to l respectively. The following statement is true (Vasin A. A., Grigorieva O. M., Tsyganov N.

[13] We are talking about such approximate methods as gradient, Newtonian and quasi-Newtonian (Sukharev A. G., Timokhov A. V., Fedorov V. B., 1985, [55]).

I., 2017, [41]).

Theorem 5. *For any line $l \in L$ the difference $\widetilde{W}(L_+ \cup L_- \cup \{l\}) - \widetilde{W}(L_+ \cup L_-)$ does not decrease by the set $L_+ \subseteq L_+(l)$ and does not increase by the set $L_- \subseteq L_-(l)$.*

Thus, if each pair of lines is in relations of complementarity or competitiveness, and these relations are established, then this theorem enables excluding from consideration some obviously suboptimal subsets of $R \subseteq L$ when solving the problem (1.7).

Denote by $\vec{Q}^0 = (Q^0_{\{i,j\}} \mid \{i,j\} \in L)$ the vector of initial transmission capacity. Consider for an arbitrary vector $\vec{Q} = (Q_{\{i,j\}} \mid \{i,j\} \in L)$, $\vec{Q} \geq \vec{Q}^0$, the initial market with modified initial transmission capacity equal to $\vec{Q}$, without the possibility of increasing transmission capacity. Let's denote by $\widetilde{q}(\vec{Q})$ the set of all possible equilibrium flows $\vec{\widetilde{q}}$ corresponding to such a market. They are determined from (1.12, 1.13) for $\vec{Q}^0 = \vec{Q}$ and $R = \emptyset$. Similarly, we introduce a set of all possible equilibrium prices $\widetilde{p}(\vec{Q})$.

Definition 6. *Let's call the vector of* <u>*directions*</u> *vector $\vec{d} = (d_{ij} \mid \{i,j\} \in L)$, for which $d_{ij} \in \{-1, 1\}$, $d_{ij} = -d_{ji}$, $\{i,j\} \in L$. In this case, $d_{ij} = 1$ corresponds to the direction from node i to node j, and $d_{ij} = -1$ - corresponds to the opposite direction.*

Definition 7. *The considered model of the market satisfies* <u>*the flow structure invariance condition*</u> *(FSIC), if there is such a directions vector $\vec{d}$, that for any $\vec{Q} \geq \vec{Q}^0$, $\vec{\widetilde{q}} \in \widetilde{q}(\vec{Q})$ the following conditions are met:*

$$\operatorname{sgn}(\widetilde{q}_{ij}) \in \{0, d_{ij}\}, \ \{i,j\} \in L,$$

where

$$\operatorname{sgn}(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0. \end{cases}$$

Thus, when performing a FSIC, with any increase in the transmission capacity for all lines, the directions of the equilibrium flows are preserved. Consider two arbitrary lines l and r, $l \neq r$. In graph $G = (N, L)$ characterizing the transport structure of the market, there is a single simple path starting with l and ending with r.

Definition 8. *Line $l = \{l_1, l_2\}$ is called* <u>*initially complementary*</u> *(respectively* <u>*initially competitive*</u>*) for the line $r = \{r_1, r_2\}$ if there are directions $d_{l_1 l_2}, d_{r_1 r_2} \in \{-1, 1\}$ such that at initial transmission capacities*

1) any equilibrium flows are consistent with these directions: $\forall\ \vec{\widetilde{q}} \in \widetilde{q}(\vec{Q}^0)$ *the inclusions* $\operatorname{sgn}(\widetilde{q}_{l_1l_2}) \in \{0, d_{l_1l_2}\}$, $\operatorname{sgn}(\widetilde{q}_{r_1r_2}) \in \{0, d_{r_1r_2}\}$ *are met;*

2) these directions are the same (respectively opposite) relative to the path connecting lines l *and* r.

If the directions of equilibrium flows are known at initial transmission capacities, then all initially complementary and initially competitive lines are easily determined for each line (figure 10).

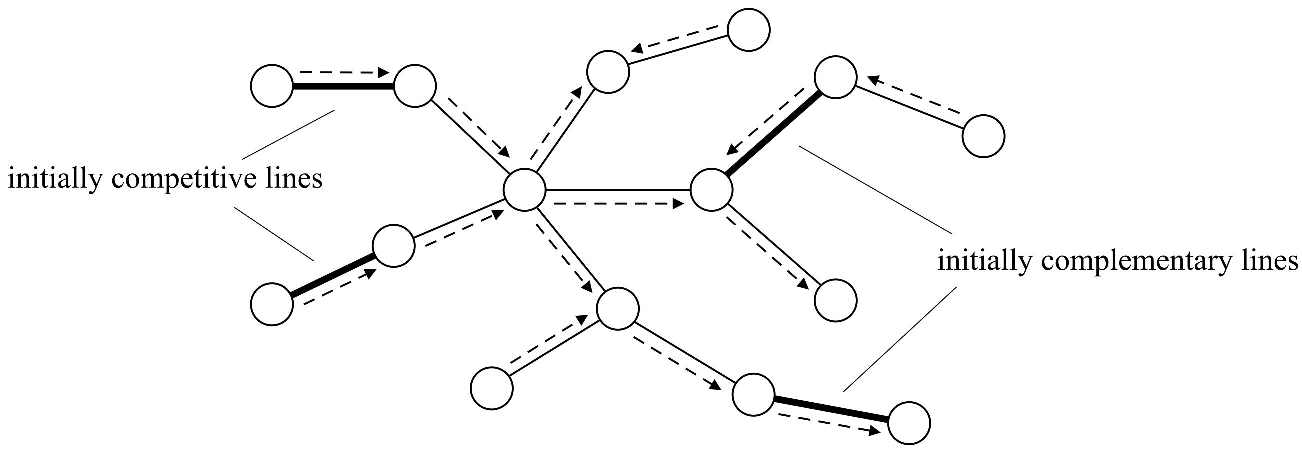

Fig. 10. Initially complementary and initially competitive lines. The arrows indicate the directions of equilibrium flows at initial transmission capacities

Denote by $\vec{Q}^\infty$ the vector of dimension $|L|$, all components of which are equal to plus infinity. Let for each line $l = \{i, j\} \in L$ the direction d_{ij} of the equilibrium flow at initial transmission capacities be uniquely determined, and $L^0_+(l)$ and $L^0_-(l)$ are sets of initially complementary and initially competitive lines to l respectively. Then the following criterion for the implementation of FSIC is valid (Vasin A. A., Grigorieva O. M., Tsyganov N. I., 2017, [41]).

Theorem 6. *The market satisfies the FSIC if and only if for any line* $l = \{i, j\} \in L$ *and for any equilibrium flows* $\vec{\widetilde{q}} \in \widetilde{q}\left(\vec{Q}^0_{L^0_+(l)} + \vec{Q}^\infty_{L^0_-(l)}\right)$ *the ratio* $\operatorname{sgn}(\widetilde{q}_{ij}) \in \{0, d_{ij}\}$ *is met, where* $\vec{Q}^0_{L^0_+(l)}$ *and* $\vec{Q}^\infty_{L^0_-(l)}$ *are the projections of vectors* $\vec{Q}^0$ *and* $\vec{Q}^\infty$ *on the subspaces* $L^0_+(l)$ *and* $L^0_-(l)$ *respectively. If this condition is met,* $L_+(l) = L^0_+(l)$, $L_-(l) = L^0_-(l)$ *for* $\forall\ l \in L$.

Theorem 6 enables establishing the relations of complementarity and competitiveness for lines. It is worth noting that if for some line $\{i, j\}$ the inequality $e^t_{\{i,j\}} > 0$ is true, then $\forall\ \vec{Q} \geq \vec{Q}^0$, $\vec{\widetilde{q}}^1, \vec{\widetilde{q}}^2 \in \widetilde{q}(\vec{Q})$ condition $\operatorname{sgn}(\widetilde{q}^1_{ij}) \cdot \operatorname{sgn}(\widetilde{q}^2_{ij}) \geq 0$ is true, i.e. two different equilibrium flows cannot have opposite directions, which simplifies the application criteria. The definition of FSIC and the invariance criterion are equivalent to those described in [41] (Vasin A. A., Grigorieva O. M., Tsyganov N. I., 2017).

Let's assume that for the market, the transport structure of which is shown in figure 11, FSIC is implemented, and the directions of equilibrium flows are marked with arrows.

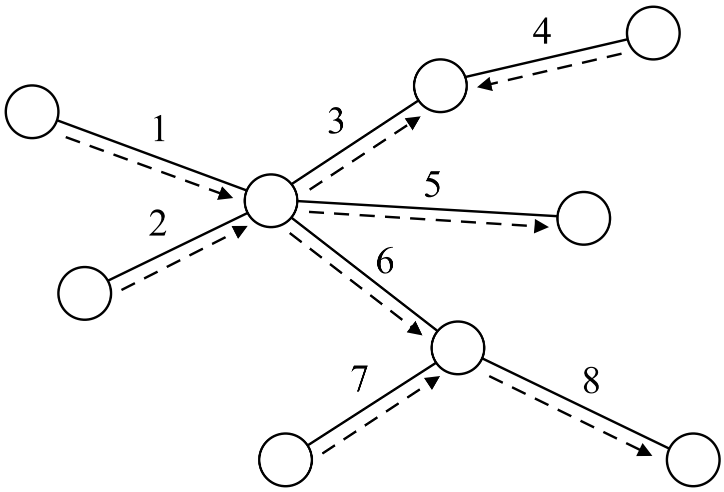

Fig. 11. An example of the transport structure of the market. The arrows indicate the directions of equilibrium flows

Example 1. Let's write out a sufficient condition under which the extension of line 1 is obviously optimal. We can find the set of complementary and competitive lines using the figure:

- $L_+(1) = \{3, 5, 6, 8\}$ is a set of lines that are complementary to 1;
- $L_-(1) = \{2, 4, 7\}$ is a set of lines that are competitive to 1.

According to theorem 5, $\forall\ M \subseteq L \setminus \{1\}$ the inequality

$$\widetilde{W}(L_-(1) \cup \{1\}) - \widetilde{W}(L_-(1)) \leq \widetilde{W}(M \cup \{1\}) - \widetilde{W}(M)$$

holds, which means that when

$$\widetilde{W}(L_-(1) \cup \{1\}) - \widetilde{W}(L_-(1)) \geq 0 \Leftrightarrow \widetilde{W}(\{2, 4, 7\}) \leq \widetilde{W}(\{1, 2, 4, 7\}) \tag{1.21}$$

$\forall\ M \subseteq L \setminus \{1\}$, the inequality

$$\widetilde{W}(M) \leq \widetilde{W}(M \cup \{1\}),$$

is true, that is, the extension of line 1 is obviously optimal. In (1.21), the values of the social welfare function are considered when expanding only lines competitive to line 1. If in such a «worst-case» for this line, with its extension, the welfare does not decrease, then it will not decrease in any other case. Similarly, a condition is written out in which, on the contrary, the absence of an extension of line 1 is obviously optimal. It already expands the lines complementary to line 1 (the «best» case), and compares the welfare values with the extension and absence of extension of line 1. This condition has the following form:

$$\widetilde{W}(\{1,3,5,6,8\}) \leq \widetilde{W}(\{3,5,6,8\}).$$

1.5. Special cases of the problem for which polynomial solving algorithms exist

Although the original problem (1.5), as proved in paragraph 1.2., is NP-hard, there are polynomial algorithms for solving it for some transport structures. Let's name by «chain» a type of market in which nodes are connected sequentially and numbered from 1 to $n = |N|$ (figure 12): $N = \{1, ..., n\}$, $L = \{\{1, 2\}, ..., \{n-1, n\}\}$.

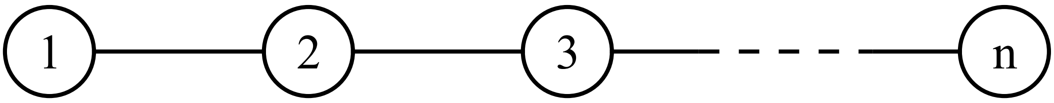

Fig. 12. Chain-type market

In the work of Vasin A. A. and Dolmatova M. S. [31] a chain-type market is considered, in which the initial transmission capacity $Q^0_{\{i,i+1\}}$ is zero, and the initial equilibrium prices $\widetilde{p}_i$ are monotonous with respect to i for $\forall\ \overrightarrow{\widetilde{p}} \in \widetilde{p}(\vec{Q}^0)$. An algorithm for solving the problem for such a market is described. It boils down to solving no more than $\frac{n(n-1)}{2}$ auxiliary problems (1.6).

In this section, the following two generalizations of the case considered in [31] are considered:

1) a chain-type market with zero initial transmission capacity;
2) a chain-type market with initial equilibrium prices monotonous with respect to i.

For each case, an algorithm of polynomial complexity is described for solving the problem of finding the optimal set of expanded lines (1.7), which the original problem (1.5) reduces to. Complexity refers to the number of auxiliary problems to be solved (1.6).

1.5.1. Chain-type market with zero initial transmission capacity

Let's consider a chain-type market in which $Q^0_l = 0$ for $\forall\ l \in L$. Such a market has the following feature: if some line $\{i, i+i\} \in L$ obviously does not expand, then the original market is divided into two isolated sub-markets with nodes $N_1 = \{1, ..., i-1\}$, $N_2 = \{i+1, ..., n\}$ and lines

$L_1 = \{\{1,2\},...,\{i-1,i\}\}$, $L_2 = \{\{i+1,i+2\},...,\{n-1,n\}\}$ respectively, therefore the problem (1.7) for the initial market is reduced to solving two similar problems of smaller dimension for the sub-markets, the maximum welfare for the source market is equal to the sum of the maximum welfare for the sub-markets.

Denote by $i_1 \leftrightarrow i_2$ an isolated sub-market with sets of nodes $\{i_1,...,i_2\}$ and lines $\{\{i_1,i_1+1\},...,\{i_2-1,i_2\}\}$, $1 \le i_1 \le i_2 \le n$. Let's denote the following values for the sub-market $i_1 \leftrightarrow i_2$: $W^*_{i_1 \leftrightarrow i_2}$ - maximum welfare, $L^*_{i_1 \leftrightarrow i_2}$ - optimal set of expanded lines.

The algorithm for solving the problem (1.7) for such a market consists in calculating the values $W^*_{i_1 \leftrightarrow i_2}$, $L^*_{i_1 \leftrightarrow i_2}$ numbered in ascending order $i_2 - i_1$: first, the sub-markets with one node are considered, then with two, etc. As a result, the desired maximum welfare and the optimal set of expandable lines for the initial market are equal, to $W^*_{1 \leftrightarrow n}$ and $L^*_{1 \leftrightarrow n}$ respectively.

The algorithm consists of the following steps.

Step 1. For $\forall\ i \in N$:

1) the value of $W^*_{i \leftrightarrow i}$ is calculated (the auxiliary problem (1.6) is solved for sub-market $i \leftrightarrow i$ with $R = \emptyset$);
2) assignment $L^*_{i \leftrightarrow i} = \emptyset$ is performed.

Step k, k=2, ..., n. For each sub-market $i_1 \leftrightarrow i_2$ such that $i_2 - i_1 = k - 1$:

1) the maximum welfare $\overline{W}_{i_1 \leftrightarrow i_2}$ is calculated for sub-market $i_1 \leftrightarrow i_2$ with the expansion of all its lines (the auxiliary problem (1.6) is solved for this sub-market with $R = \overline{L}_{i_1 \leftrightarrow i_2} \equiv \{\{i_1,i_1+1\},...,\{i_2-1,i_2\}\}$);
2) for each node $j \in [i_1, i_2-1]$ the maximum welfare $W^*_{i_1 \leftrightarrow i_2, j}$ and the optimal set of expandable lines $L^*_{i_1 \leftrightarrow i_2, j}$ for sub-market $i_1 \leftrightarrow i_2$ are calculated with an obvious non-expansion of the line $\{j, j+1\}$, in this case sub-market $i_1 \leftrightarrow i_2$ is divided into two isolated sub-markets $i_1 \leftrightarrow j$ and $j+1 \leftrightarrow i_2$:

$$W^*_{i_1 \leftrightarrow i_2, j} = W^*_{i_1 \leftrightarrow j} + W^*_{j+1 \leftrightarrow i_2},\ L^*_{i_1 \leftrightarrow i_2, j} = L^*_{i_1 \leftrightarrow j} \cup L^*_{j+1 \leftrightarrow i_2};$$

3) for the obtained pairs

$$A_{i_1 \leftrightarrow i_2} = \left\{ \left(\overline{W}_{i_1 \leftrightarrow i_2}, \overline{L}_{i_1 \leftrightarrow i_2}\right), \left(W^*_{i_1 \leftrightarrow i_2, i_1}, L^*_{i_1 \leftrightarrow i_2, i_1}\right), ..., \left(W^*_{i_1 \leftrightarrow i_2, i_2-1}, L^*_{i_1 \leftrightarrow i_2, i_2-1}\right) \right\}$$

the one for which welfare is maximized is chosen:

$$W^*_{i_1 \leftrightarrow i_2} = \max_{(W,L') \in A_{i_1 \leftrightarrow i_2}} W,\ L^*_{i_1 \leftrightarrow i_2} \in \left\{ L' \mid (W^*_{i_1 \leftrightarrow i_2}, L') \in A_{i_1 \leftrightarrow i_2} \right\}.$$

As a result, $L^* = L^*_{1\leftrightarrow n}$ is the optimal set of expanded lines for the entire market, and $W^* = W^*_{1\leftrightarrow n}$ is maximum public welfare. The correctness of the algorithm follows from the fact that when calculating the $W^*_{i_1\leftrightarrow i_2}$ considered $i_2 - i_1 + 1$ cases in paragraphs 1 and 2 of step k cover all possible variants of sets of expandable lines.

Theorem 7. *The complexity of this algorithm in terms of the number of auxiliary problems to be solved (1.6) is equal to* $O\left(|N|^2\right)$.

Proof. *Indeed, the number of auxiliary problems (1.6) to be solved is the same as the number of different sub-markets* $i_1 \leftrightarrow i_2$, $1 \le i_1 \le i_2 \le n$, *and is equal to* $C_n^2 = \frac{n(n-1)}{2}$. ■

The complexity of the remaining operations[14] of the algorithm is $O\left(|N|^4\right)$. The complexity can be lowered to $O\left(|N|^3\right)$ if

1) the calculation of the set $L^*_{i_1\leftrightarrow i_2,j}$ in substep 2 of step k (k=2, ..., n) for fixed i_1, i_2, j is eliminated;
2) saving the node j in substep 2 of step k (k=2, ..., n) is added as the second element of the pair instead of $L^*_{i_1\leftrightarrow i_2,j}$;
3) in substep 3 of step k (k=2, ..., n) after finding $W^*_{i_1\leftrightarrow i_2}$ the calculation of the set $L^*_{i_1\leftrightarrow i_2}$ is added using the saved node of the selected pair.

1.5.2. Chain-type market with monotonous initial equilibrium prices

Let's consider a chain-type market in which the initial equilibrium prices $\widetilde{p}_i$ are monotonic with respect to i for $\forall\ \overrightarrow{\widetilde{p}} \in \widetilde{p}(\vec{Q}^0)$. According to the dissertation of Dolmatova M. S. [32], FSIC is performed for it, and all equilibrium flows are unidirectional (figure 13) and go towards increasing the equilibrium price, which means that any two lines $l, r \in L$, $l \neq r$, are mutually complementary.

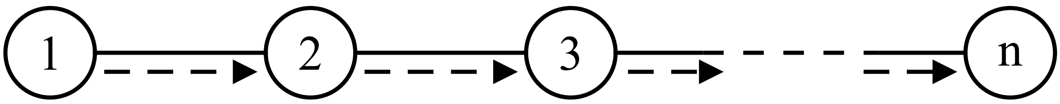

Fig. 13. A chain-type market with unidirectional flows

Without limiting generality, we assume that the equilibrium flows are directed from node 1 to node n. For a fixed set of expandable lines $R \subseteq L$, we call a line $\{i, i+1\} \in L \setminus R$ saturated if any equilibrium flow $\widetilde{q}_{i,i+1}$ from node i to node $i+1$ coincides with the initial transmission capacity

[14] Without taking into account the operations necessary to solve auxiliary problems.

$Q^0_{\{i,i+1\}}$, i.e. the restriction by the amount of the stream is active. Let's denote by $L^{full}(R)$ the set of all saturated lines. This market satisfies the following property: if $l \in L^{full}(R)$, then $l \in L^{full}(R')$ for any R', $R \subseteq R' \subseteq L \setminus \{l\}$. Thus, the saturation property is preserved when other lines are expanded.

It follows from this property that if, for a certain set of obviously expanded lines R, the obviously non-expanded line $\{i, i+1\}$ belongs to $L^{full}(R)$, then this line is obviously saturated, which means that regardless of the specific set of expanded lines for any equilibrium flows $\overrightarrow{\widetilde{q}}$ the condition $\widetilde{q}_{i,i+1} = Q^0_{\{i,i+1\}}$ is met, therefore the initial market is divided into two pseudo-isolated submarkets with nodes $N_1 = \{1, ..., i-1\}$, $N_2 = \{i+1, ..., n\}$ and lines $L_1 = \{\{1,2\}, ..., \{i-1, i\}\}$, $L_2 = \{\{i+1, i+2\}, ..., \{n-1, n\}\}$, respectively, which enables reducing the problem (1.7) for the initial market to solving two similar problems of smaller dimension for the sub-markets taking into account the boundary condition $q_{i,i+1} = Q^0_{\{i,i+1\}}$, while the maximum welfare for the initial market is equal to the sum of the maximum welfare for the sub-markets minus transportation costs for the boundary line $E_{i,i+1}(Q^0_{\{i,i+1\}})$.

For convenience, we introduce fictitious lines $\{0,1\}$, $\{n, n+1\}$ with initial transmission capacity $Q^0_{\{0,1\}} = Q^0_{\{n,n+1\}} = 0$, as well as fictitious streams $q_{0,1} = q_{n,n+1} = 0$. Let's denote through $i_1 \to i_2$, $1 \leq i_1 \leq i_2 \leq n$, a pseudo-isolated sub-market with sets of nodes $\{i_1, ..., i_2\}$ and inner lines $\{\{i_1, i_1+1\}, ..., \{i_2-1, i_2\}\}$, for which boundary lines $\{i_1-1, i_1\}$ and $\{i_2, i_2+1\}$ are obviously saturated with flows $Q^0_{\{i_1-1,i_1\}}$ and $Q^0_{\{i_2,i_2+1\}}$ respectively (figure 14).

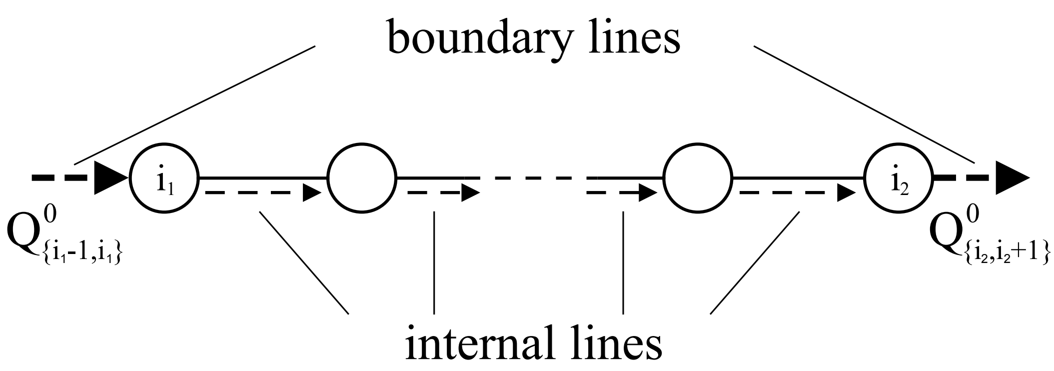

Fig. 14. Pseudo-isolated sub-market $i_1 \to i_2$

We denote the following values for the sub-market $i_1 \to i_2$: $W^*_{i_1 \to i_2}$ is maximum welfare (excluding transport costs for boundary lines), $L^*_{i_1 \to i_2}$ is the optimal set of expandable internal lines.

An algorithm for finding the optimal set of expandable lines (i.e. solving the problem (1.7)) for this market consists in the sequential calculation of values $W^*_{i_1 \to i_2}$, $L^*_{i_1 \to i_2}$ in ascending order $i_2 - i_1$: first, the sub-markets with one node are considered, then with two, etc. As a result, the desired maximum welfare and the optimal set of expandable lines for the initial market are equal to $W^*_{1\to n}$ and $L^*_{1\to n}$ respectively.

Let's denote for the sub-market $i_1 \to i_2$ by $L_{i_1\to i_2} = \{\{i_1, i_1+1\}, ..., \{i_2-1, i_2\}\}$ the set of internal lines, and through $L^{full}_{i_1\to i_2}(R)$ and $\widetilde{W}_{i_1\to i_2}(R)$ the set of all saturated internal lines and maximum welfare, respectively, for a fixed set of expandable lines $R \subseteq L_{i_1\to i_2}$. It is worth noting that if the condition is satisfied for a certain set of obviously expanded lines $R \subseteq L_{i_1\to i_2}$, condition $L^{full}_{i_1\to i_2}(R) = \emptyset$ is met, then R is the optimal set of expanded internal lines. This is obvious, because in the absence of saturated lines, there is no need to expand.

The algorithm consists of the following steps.

Step 1. For $\forall\ i \in N$:

1) the value $W^*_{i\to i}$ is calculated (the auxiliary problem (1.6) is solved for the sub-market $i \to i$ with $R = \emptyset$ and boundary conditions $q_{i-1,i} = Q^0_{\{i-1,i\}}$, $q_{i,i+1} = Q^0_{\{i,i+1\}}$);
2) assignment $L^*_{i\to i} = \emptyset$ is performed.

Step k, k=2, ..., n. For each sub-market $i_1 \to i_2$ such that $i_2 - i_1 = k - 1$:

1) assignments $R_{i_1\to i_2} = \emptyset$, $A_{i_1\to i_2} = \emptyset$ are performed (the variable $R_{i_1\to i_2}$ defines a set of obviously extensible inner lines, $A_{i_1\to i_2}$ is a set of pairs of the form (W, L'), where $W = \widetilde{W}_{i_1\to i_2}(L')$, $L' \subseteq L_{i_1\to i_2}$);
2) the set of saturated lines $L^{full}_{i_1\to i_2}(R_{i_1\to i_2})$ and the social welfare $\widetilde{W}_{i_1\to i_2}(R_{i_1\to i_2})$ are calculated (the auxiliary problem (1.6) is solved for sub-market $i_1 \to i_2$ with $R = R_{i_1\to i_2}$ and boundary conditions $q_{i_1-1,i_1} = Q^0_{\{i_1-1,i_1\}}$, $q_{i_2,i_2+1} = Q^0_{\{i_2,i_2+1\}}$);
3) if $L^{full}_{i_1\to i_2}(R_{i_1\to i_2}) \neq \emptyset$, then the saturated line $\{j, j+1\} \in L^{full}_{i_1\to i_2}(R_{i_1\to i_2})$ is randomly selected, after which two cases are considered sequentially:
   3.1) the line $\{j, j+1\}$ does not expand: in this case, sub-market $i_1 \to i_2$ is divided into $i_1 \to j$ and $j+1 \to i_2$, so a pair $\left(W^*_{i_1\to j} - E_{j,j+1}(Q^0_{\{j,j+1\}}) + W^*_{j+1\to i_2}, L^*_{i_1\to j} \cup L^*_{j+1\to i_2}\right)$ is added to $A_{i_1\to i_2}$;
   3.2) the line $\{j, j+1\}$ is expanded: the line$\{j, j+1\}$ is added to the set $R_{i_1\to i_2}$; a return to substep 2 occurs;

   if $L^{full}_{i_1\to i_2}(R_{i_1\to i_2}) = \emptyset$, then the pair $\left(\widetilde{W}_{i_1\to i_2}(R_{i_1\to i_2}), R_{i_1\to i_2}\right)$ is added to $A_{i_1\to i_2}$;
4) for the found pairs $A_{i_1\to i_2}$ the one for which the welfare is maximum is selected:

$$W^*_{i_1\to i_2} = \max_{(W,L')\in A_{i_1\to i_2}} W,\ L^*_{i_1\to i_2} \in \left\{L' \mid (W^*_{i_1\to i_2}, L') \in A_{i_1\to i_2}\right\}.$$

As a result, $L^* = L^*_{1\to n}$ is the optimal set of expanded lines for the entire market, and $W^* = W^*_{1\to n}$ is maximum public welfare. The correctness of the algorithm follows from the properties of this type of market described above.

Theorem 8. *The complexity of this algorithm in terms of the number of auxiliary problems to be solved (1.6) is equal to* $O\left(|N|^3\right)$*.*

Proof. *Indeed, the number of sub-markets considered* $i_1 \to i_2$*,* $1 \le i_1 \le i_2 \le n$*, is equal to* $C_n^2 = \frac{n(n-1)}{2}$*, and for each sub-market* $i_1 \to i_2$*, no more than* $i_2 - i_1 + 1 \le |N|$ *auxiliary problems are solved (1.6).* ■

The complexity of the remaining operations[15] of the algorithm is $O\left(|N|^4\right)$.

1.6. Solving the problem when the flow structure invariance condition is met

In this section, we consider the markets for which FSIC is met, i.e. the directions of equilibrium flows are preserved. In such markets, any pair of lines is in a relationship of complementarity or competitiveness, which can be established based on known directions. Consideration of this case requires special attention, since in practice, flows often have known directions for existing and potential lines and go from producing nodes (for example, gas or oil fields) to consuming ones (industrial centers, cities, etc.).

Algorithms for solving the problem of finding the optimal set of expandable lines (1.7) are described below for various transport structures. The idea of all the presented algorithms is based on the application of theorem 5. These algorithms are not polynomial, therefore, to assess their effectiveness, an indicator such as the average statistical complexity is used - the average number of solved auxiliary tasks (1.6) for a set of randomly generated initial tasks. We assume that for each line $l \in L$, the direction of the equilibrium flow and the set of complementary $L_+(l)$ and competitive $L_-(l)$ lines are known.

1.6.1. Chain-type market

Let's consider a chain-type market with sets of nodes $N = \{1, ..., n\}$ and lines $L = \{\{1,2\}, ..., \{n-1, n\}\}$. We divide the set of lines into two subsets: L_1 with flow directions from

[15] Without taking into account the operations necessary to solve auxiliary problems.

node 1 and L_2 with flow directions to node 1 (figure 15). In this case, two arbitrary lines are mutually complementary if they belong to the same subset, and mutually competitive otherwise.

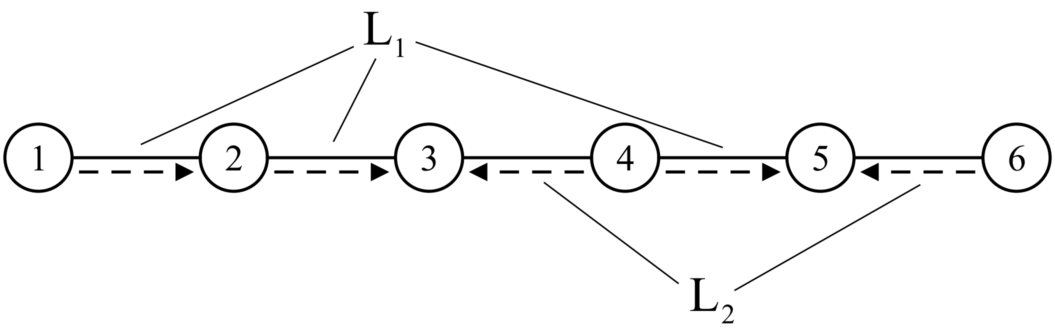

Fig. 15. An example of a chain-type market with $L_1 = \{\{1,2\},\{2,3\},\{4,5\}\}$, $L_2 = \{\{3,4\},\{5,6\}\}$

Let's proceed to the construction of an algorithm for finding the optimal set of expanded lines L^*. For any $R \subseteq L$, we denote $L_1(R) = L_1 \cap R$, $L_2(R) = L_2 \cap R$. Then the desired set L^* is represented as a union of sets $L_1^* = L_1(L^*)$ and $L_2^* = L_2(L^*)$. The algorithm works with *lower* $L_1^{\min}$ and *upper* $L_1^{\max}$ *estimates* of the set L_1^*, as well as the lower $L_2^{\min}$ and the upper $L_2^{\max}$ estimates of the set L_2^* and is based on the sequential application of the steps of two types. In the first type of steps, attempts are made to expand the set $L_1^{\min}$ and narrow the set $L_2^{\max}$. In the second type of steps, attempts are made to narrow the set $L_1^{\max}$ and expand the set $L_2^{\min}$. Let's call a line *included*, if it belongs to one of the current lower estimates, and *excluded*, if it does not belong to any of the current upper estimates. Let's call a line *defined*, if it is included or excluded. If for the current estimates for a step of the first type, the number of lines defined on it is no more than for a step of the second type, then a step of the first type is performed. Otherwise, the second type of step is performed. As a result, the optimal set L^* is searched for in the form of $L_1^* \cup L_2^*$, where $L_1^{\min} \subseteq L_1^* \subseteq L_1^{\max}$, $L_2^{\min} \subseteq L_2^* \subseteq L_2^{\max}$.

The validity of the following statement follows from theorem 5.

Theorem 9. *Let $L_1^{\min}$ and $L_1^{\max}$ be the current lower and upper bounds of the set L_1^*, and $L_2^{\min}$ and $L_2^{\max}$ - the current lower and upper estimates of the set L_2^*. Let $S_1 \subseteq L_1^{\max} \setminus L_1^{\min}$, $S_2 \subseteq L_2^{\max} \setminus L_2^{\min}$, $S = S_1 \cup S_2$.*

1) If the inequality

$$\widetilde{W}\left((L_1^{\min} \cup S_1) \cup (L_2^{\max} \setminus S_2)\right) \geq \widetilde{W}\left(L_1^{\min} \cup L_2^{\max}\right)$$

is met and for each nonempty set $R \subset S$

$$\widetilde{W}\left((L_1^{\min} \cup L_1(R)) \cup (L_2^{\max} \setminus L_2(R))\right) < \widetilde{W}\left(L_1^{\min} \cup L_2^{\max}\right),$$

then the set S_1 *can be added to the lower estimate* $L_1^{\min}$*, and the set* S_2 *can be subtracted from the upper estimate* $L_2^{\max}$*:* $L_{1,r}^{\min} = L_1^{\min} \cup S_1$*,* $L_{2,r}^{\max} = L_2^{\max} \setminus S_2$*, where* $L_{1,r}^{\min}$*,* $L_{2,r}^{\max}$ *are adjusted estimates.*

2) If the inequality

$$\widetilde{W}\left((L_1^{\max} \setminus S_1) \cup (L_2^{\min} \cup S_2)\right) \geq \widetilde{W}\left(L_1^{\max} \cup L_2^{\min}\right)$$

is met and for each nonempty set $R \subset S$

$$\widetilde{W}\left((L_1^{\max} \setminus L_1(R)) \cup (L_2^{\min} \cup L_2(R))\right) < \widetilde{W}\left(L_1^{\max} \cup L_2^{\min}\right),$$

then the set S_1 *can be subtracted from the upper estimate* $L_1^{\max}$*, and the set* S_2 *can be added to the lower estimate* $L_2^{\min}$*:* $L_{1,r}^{\max} = L_1^{\max} \setminus S_1$*,* $L_{2,r}^{\min} = L_2^{\min} \cup S_2$*, where* $L_{1,r}^{\max}$*,* $L_{2,r}^{\min}$ *are adjusted estimates.*

Proof. *Let's prove point 1. Point 2 is proved similarly. Let the specified inequalities be satisfied. It is enough to show that for any* $S_1' \subseteq L_1^{\max} \setminus L_1^{\min}$*,* $S_2' \subseteq L_2^{\max} \setminus L_2^{\min}$ *the following ratio is true:*

$$\widetilde{W}\left(((L_1^{\min} \cup S_1') \cup S_1) \cup ((L_2^{\max} \setminus S_2') \setminus S_2)\right) \geq \widetilde{W}\left((L_1^{\min} \cup S_1') \cup (L_2^{\max} \setminus S_2')\right). \qquad (1.22)$$

Let's denote $S_1'' = S_1' \setminus S_1$*,* $S_2'' = S_2' \setminus S_2$*. Let's number the elements of the following sets:* $S_1 \setminus S_1' = \{l_1^1, ..., l_1^{m_1}\}$*,* $S_2 \setminus S_2' = \{l_2^1, ..., l_2^{m_2}\}$*. Let's consider the difference of the expressions from (1.22):*

$$\begin{aligned}
&\widetilde{W}\left(((L_1^{\min} \cup S_1') \cup S_1) \cup ((L_2^{\max} \setminus S_2') \setminus S_2)\right) - \widetilde{W}\left((L_1^{\min} \cup S_1') \cup (L_2^{\max} \setminus S_2')\right) = \\
&= \widetilde{W}\left(((L_1^{\min} \cup S_1) \cup S_1'') \cup ((L_2^{\max} \setminus S_2) \setminus S_2'')\right) - \\
&- \widetilde{W}\left(((L_1^{\min} \cup (S_1 \cap S_1')) \cup S_1'') \cup ((L_2^{\max} \setminus (S_2 \cap S_2')) \setminus S_2'')\right) = \\
&= \widetilde{W}\left(((L_1^{\min} \cup S_1) \cup S_1'') \cup ((L_2^{\max} \setminus S_2) \setminus S_2'')\right) \pm \\
&\pm \sum_{k=1}^{m_1} \widetilde{W}\left(((L_1^{\min} \cup (S_1 \setminus \{l_1^1, ..., l_1^k\})) \cup S_1'') \cup ((L_2^{\max} \setminus S_2) \setminus S_2'')\right) \pm \\
&\pm \sum_{k=1}^{m_2-1} \widetilde{W}\left(((L_1^{\min} \cup (S_1 \cap S_1')) \cup S_1'') \cup ((L_2^{\max} \setminus (S_2 \setminus \{l_2^1, ..., l_2^k\})) \setminus S_2'')\right) \pm \\
&- \widetilde{W}\left(((L_1^{\min} \cup (S_1 \cap S_1')) \cup S_1'') \cup ((L_2^{\max} \setminus (S_2 \cap S_2')) \setminus S_2'')\right).
\end{aligned}$$

By rearranging the terms in the last expression, applying the theorem 5 $m_1 + m_2$ *times and making reductions, it can be proved that the last expression is not less than*

$$\widetilde{W}\left((L_1^{\min} \cup S_1) \cup (L_2^{\max} \setminus S_2)\right) - \widetilde{W}\left((L_1^{\min} \cup (S_1 \cap S_1')) \cup (L_2^{\max} \setminus (S_2 \cap S_2'))\right) \geq 0,$$

from which validity of (1.22) follows. The last transition takes into account the inclusions $S_1 \cap S_1' \subseteq S_1$, $S_2 \cap S_2' \subseteq S_2$ *and the initial inequalities.* ■

1.6.1.1. Algorithm

Let's describe an **algorithm** for finding the optimal set of expanded lines. It uses the following variables: $L_1^{\min}$, $L_1^{\max}$, $L_2^{\min}$, $L_2^{\max}$ - the current lower and upper estimates for sets L_1^* and L_2^*; k_1 and k_2 - the number of defined lines for the steps of the first and second type respectively; $\overline{L}$ is the set of lines not yet defined; $T \in \{1, 2\}$ is the type of the current step; s is the step number.

1) Assignments are performed: $\overline{L} = L$, $L_1^{\min} = L_2^{\min} = \emptyset$, $L_1^{\max} = L_1$, $L_2^{\max} = L_2$, $k_1 = k_2 = 1$, $s = 1$.

2) Assignment is performed: $T = \begin{cases} 1, & k_1 \leq k_2, \\ 2, & k_1 > k_2. \end{cases}$

3) If $s = n - 1$, then the optimal set is chosen according to

$$L^* \in \underset{R \in \{L_1^{\min} \cup L_2^{\max}, L_1^{\max} \cup L_2^{\min}\}}{\text{Arg max}} \widetilde{W}(R),$$

after that the algorithm shuts down.

4) If $T = 1$, then:

4.1) all sets S for which $S \subseteq \overline{L}$, $|S| = k_1$, are considered sequentially; for each such set, the inequality

$$\widetilde{W}\left((L_1^{\min} \cup L_1(S)) \cup (L_2^{\max} \setminus L_2(S))\right) \geq \widetilde{W}\left(L_1^{\min} \cup L_2^{\max}\right)$$

is checked, if successful, the following assignments are performed without considering the remaining sets: $L_1^{\min} = L_1^{\min} \cup L_1(S)$, $L_2^{\max} = L_2^{\max} \setminus L_2(S)$, $\overline{L} = \overline{L} \setminus S$, $k_1 = 1$, $s = s + 1$, then a return to step 2 occurs;

4.2) assignments $k_1 = k_1 + 1$, $s = s + 1$ are performed, after which a return to point 2 occurs.

If $T = 2$, then:

4.1) all sets S, for which $S \subseteq \overline{L}$, $|S| = k_2$, are considered sequentially; for each such set, the inequality

$$\widetilde{W}\left((L_1^{\max} \setminus L_1(S)) \cup (L_2^{\min} \cup L_2(S))\right) \geq \widetilde{W}\left(L_1^{\max} \cup L_2^{\min}\right)$$

is checked, if successful, the following assignments are performed without considering the remaining sets: $L_1^{\max} = L_1^{\max} \setminus L_1(S)$, $L_2^{\min} = L_2^{\min} \cup L_2(S)$, $\overline{L} = \overline{L} \setminus S$, $k_2 = 1$,

$s = s + 1$, then a return to step 2 occurs;

4.2) assignments $k_2 = k_2 + 1$, $s = s + 1$ are performed, after which a return to point 2 occurs.

Note that if for the set S the inequality $\widetilde{W}\left((L_1^{\min} \cup L_1(S)) \cup (L_2^{\max} \setminus L_2(S))\right) \geq \widetilde{W}\left(L_1^{\min} \cup L_2^{\max}\right)$ holds, then, according to theorem 9, the set $L_1(S)$ can be added to the lower estimate $L_1^{\min}$, and the set $L_2(S)$ can be subtracted from the upper estimate $L_2^{\max}$. Indeed, every nonempty set $R \subset S$ has already been considered in the previous steps and satisfies the inequality $\widetilde{W}\left((L_1^{\min} \cup L_1(R)) \cup (L_2^{\max} \setminus L_2(R))\right) < \widetilde{W}\left(L_1^{\min} \cup L_2^{\max}\right)$ (otherwise, the set $L_1(R)$ would have already been added to the lower estimate $L_1^{\min}$, and the set $L_2(R)$ would have already been subtracted from the upper estimate $L_2^{\max}$). For the second type of steps, similar reasoning is true. This implies the correctness of the algorithm.

1.6.1.2. Estimation of the average complexity of the algorithm

The described algorithm enables significantly reducing the number of auxiliary problems (1.6) to be solved when solving the problem (1.7). To assess the dependence of the number of auxiliary problems to be solved on the number of nodes n a computational experiment was conducted, during which a set of initial problems was randomly generated for a different number of nodes, after which each problem was solved by the developed algorithm.

Let's describe the market model used in the experiment. Nodes are characterized by piecewise linear supply and demand functions of the following types:

$$S_i(p_i) = \begin{cases} \frac{1}{2}c_i \cdot p_i, & p_i \leq 2\frac{d_i^f}{c_i}, \\ -d_i^f + c_i \cdot p_i, & p_i > 2\frac{d_i^f}{c_i}, \end{cases} \qquad D_i(p_i) = \begin{cases} d_i^f - \frac{1}{2}c_i \cdot p_i, & p_i \leq 2\frac{d_i^f}{c_i}, \\ 0, & p_i > 2\frac{d_i^f}{c_i}. \end{cases} \tag{1.23}$$

These functions correspond to the linear functions of net supply $\Delta S_i(p_i) \equiv S_i(p_i) - D_i(p_i) = -d_i^f + c_i \cdot p_i$, $i \in N$. For each line $l \in L$, the function of variable costs for increasing the transmission capacity is quadratic and is characterized by the parameter e_l^q:

$$E_l^v(\Delta Q_l) = e_l^q \cdot \Delta Q_l^2. \tag{1.24}$$

Let's denote by p_i^0 the equilibrium price in node $i \in N$ in the case of its isolation (with zero transmission capacity), and by Δp_i^0 the difference between the equilibrium prices for isolated nodes i and $i+1$: $\Delta p_i^0 = p_{i+1}^0 - p_i^0$, $\{i, i+1\} \in L$. Let $p_{\min}^0 = \min\limits_{i \in N} p_i^0$. Then the initial market is uniquely determined by the following parameters: n, $p^0_{\min}$, d^f_i $(i \in N)$, Δp^0_i $(i \in \{1, ..., n-1\})$, Q^0_l, e^t_l, e^q_l, E^f_l $(l \in L)$. In this case, the equilibrium prices p^0_i, $i \in N$, are found from the system

$$\begin{cases} p^0_{i+1} - p^0_i = \Delta p^0_i, \ \{i, i+1\} \in L, \\ \min\limits_{i \in N} p^0_i = p^0_{\min}, \end{cases}$$

and $c_i = \dfrac{d^f_i}{p^0_i}$, $i \in N$ (we assume that $p^0_{\min} > 0$, $d^f_i > 0 \ \forall \ i \in N$).

Throughout the conducted experiment, the parameters $p^0_{\min}$, d^f_i, $|\Delta p^0_i|$, e^t_l, e^q_l, E^f_l, characterizing the problem were randomly generated in accordance with a uniform distribution (table 1), the initial transmission capacities Q^0_l was taken to be 0, and for any line $\{i, i+1\} \in L$ the sign of Δp^0_{i+1} coincided with the sign of Δp^0_i with a probability of 0.9, at the same time $\Delta p^0_1 > 0$.

Table 1. Parameters of probability distributions of quantities $p^0_{\min}$, d^f_i, $|\Delta p^0_i|$, e^t_l, e^q_l, E^f_l for a chain-type market

Model parameter	Minimum value	Maximum value
$p^0_{\min}$	0 (not including 0)	10
d^f_i	10	20
$\|\Delta p^0_i\|$	0 (not including 0)	10
e^t_l	0	4
e^q_l	0	4
E^f_l	0	4

In order for the market to satisfy the FSIC, for each line $\{i, i+1\} \in L$ the modified functions of marginal transport costs $e^{nex,m}_{i,i+1}(q_{i,i+1})$ and $e^{ex,m}_{i,i+1}(q_{i,i+1})$ were used. If $\operatorname{sgn}(\Delta p^0_i) > 0$, then

$$e^{nex,m}_{i,i+1}(q_{i,i+1}) = \begin{cases} \left(-\infty, \max\left(e^{nex}_{i,i+1}(0)\right)\right], & q_{i,i+1} = 0, \\ e^{nex}_{i,i+1}(q_{i,i+1}), & q_{i,i+1} > 0; \end{cases} \tag{1.25}$$

$$e^{ex,m}_{i,i+1}(q_{i,i+1}) = \begin{cases} \left(-\infty, \max\left(e^{ex}_{i,i+1}(0)\right)\right], & q_{i,i+1} = 0, \\ e^{ex}_{i,i+1}(q_{i,i+1}), & q_{i,i+1} > 0. \end{cases} \tag{1.26}$$

If $\operatorname{sgn}(\Delta p^0_i) < 0$, then

$$e^{nex,m}_{i,i+1}(q_{i,i+1}) = \begin{cases} e^{nex}_{i,i+1}(q_{i,i+1}), & q_{i,i+1} < 0, \\ \left[\min\left(e^{nex}_{i,i+1}(0)\right), +\infty\right), & q_{i,i+1} = 0; \end{cases} \tag{1.27}$$

$$e^{ex,m}_{i,i+1}(q_{i,i+1}) = \begin{cases} e^{ex}_{i,i+1}(q_{i,i+1}), & q_{i,i+1} < 0, \\ \left[\min\left(e^{ex}_{i,i+1}(0)\right), +\infty\right), & q_{i,i+1} = 0. \end{cases} \quad (1.28)$$

The modified functions (1.25-1.28) have limited definition areas, preventing the flow from going in the wrong direction. It is worth noting that this modification does not in any way violate the validity of the previously obtained results and the correctness of the developed algorithms.

The number of nodes n varied from 1 to 65. For each n 1 000 problems were generated, each of which was solved by the described algorithm. Figure 16 shows the obtained dependence of the number of auxiliary problems (1.6) to be solved on the number of nodes n.

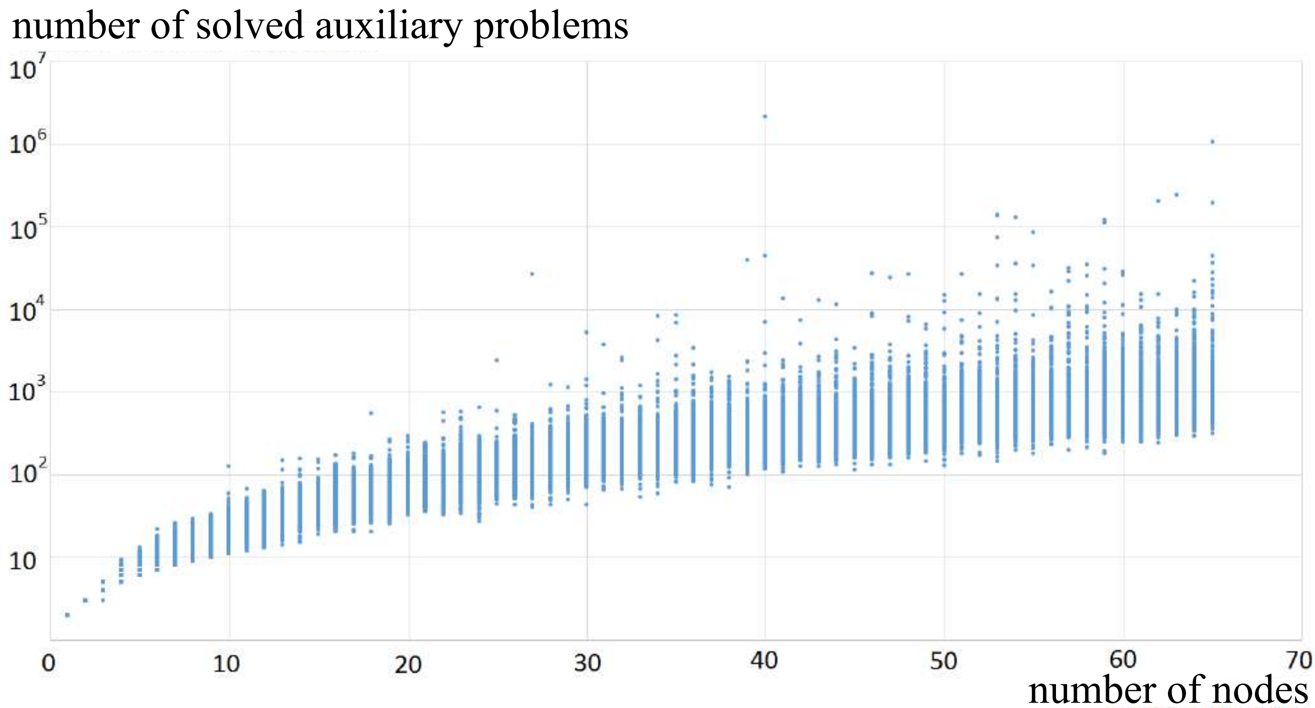

Fig. 16. The results of a numerical experiment for a chain-type market. Each point corresponds to a solved problem

Let's denote the average number of auxiliary problems solved by y_{av}. The resulting dependence of y_{av} on n is shown in 17. The following approximation of the average number of solved auxiliary problems is obtained by the least squares method: $\overline{y}_{av}(n) = 0.251n^2 + 1.771$ (figure 18). The corresponding coefficient of determination R^2 is 0.6357.

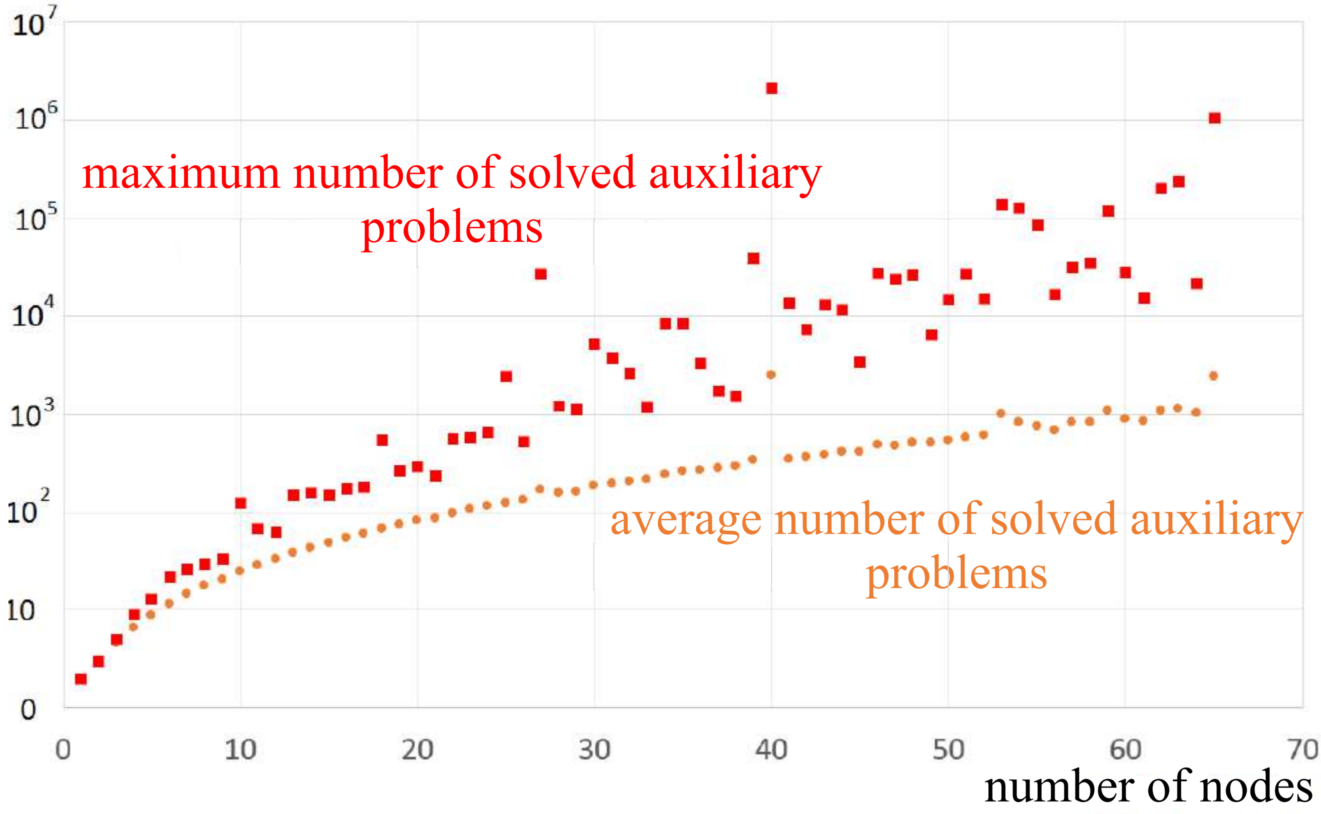

Fig. 17. Average (bottom) and maximum (top) numbers of solved auxiliary problems for the chain-type market

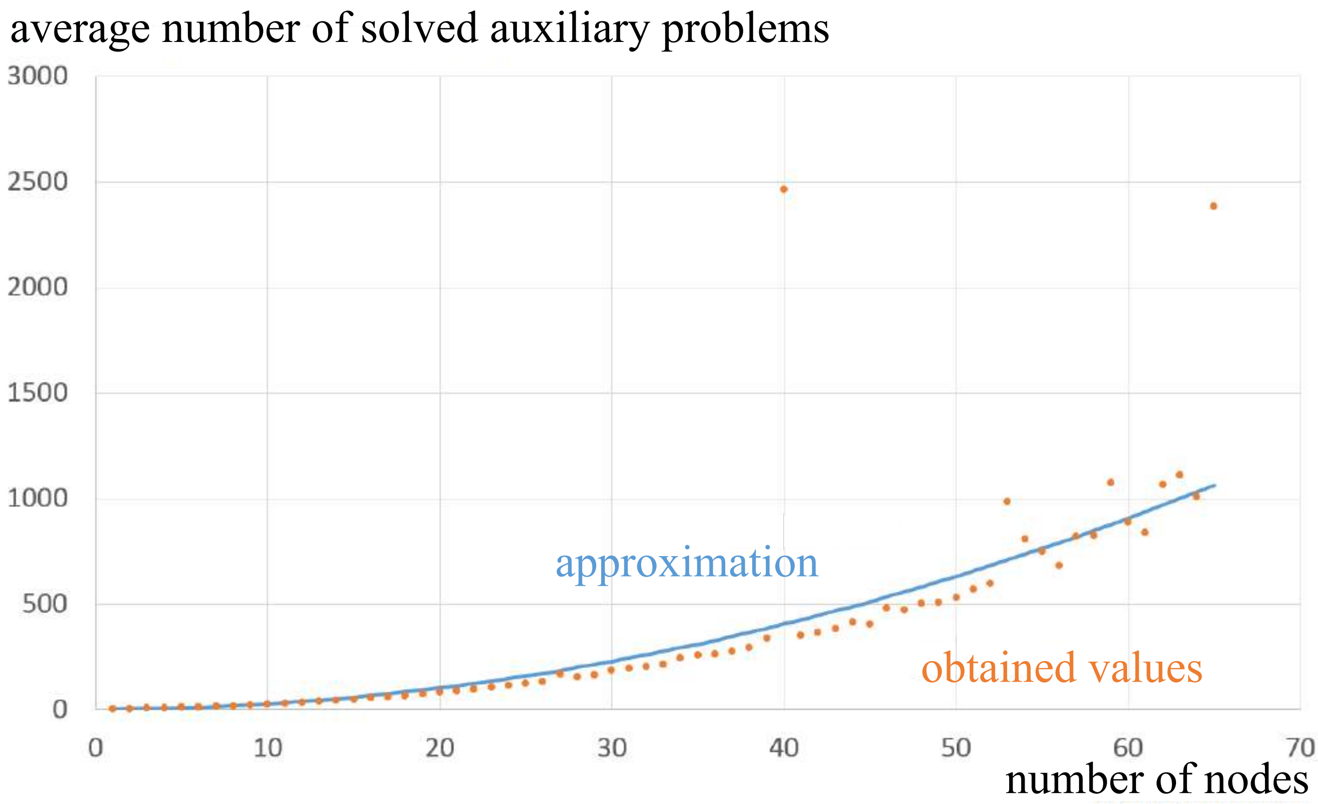

Fig. 18. Approximation of the average number of solved auxiliary problems for a chain-type market

Thus, for the conducted computational experiment, the average number of solved auxiliary problems (1.6) is well approximated by a quadratic function depending on the number of nodes. This suggests that the initial task (1.5) for this type of market can be solved in a reasonable time, although, as previously proved, it is NP-hard in the general case.

1.6.2. Star-type market

Let's consider a market of the «star» type, which consists of a central node 0 and a set of nodes adjacent to it. We divide the set of lines L into subsets L_1 with directions to the central node and L_2 with directions from the central node (figure 19). In this case, two arbitrary lines are mutually complementary if they belong to different subsets, and mutually competitive otherwise. Let's denote by $N_1 = \{i \in N \mid \{i, 0\} \in L_1\}$ the set of producing nodes, and through $N_2 = \{i \in N \mid \{0, i\} \in L_2\}$ - the set of consuming nodes.

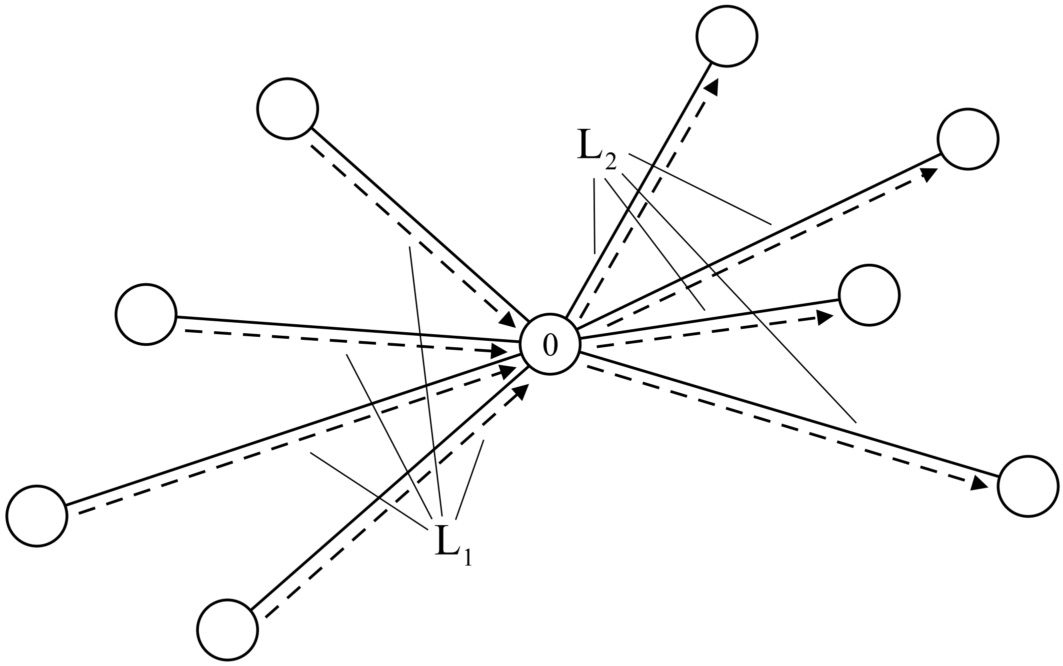

Fig. 19. An example of a star-type market

Let's proceed to the construction of an algorithm for finding the optimal set of expanded lines L^*. The algorithm works with _lower_ $L^{\min}$ and _upper_ $L^{\max}$ _estimates_ of the set L^*. Just as for the chain-type market, attempts are being made to expand the lower and narrow the upper estimates. Let's call a line _included_, if it belongs to the current lower estimate, and _excluded_, if it does not belong to the current upper estimate. Let's call a line _defined_, if it is included or excluded. As a result, the optimal set L^* satisfies the condition $L^{\min} \subseteq L^* \subseteq L^{\max}$.

Theorem 10. *Let $L^{\min}$ and $L^{\max}$ be the current lower and upper estimates, and $l \in L^{\max} \setminus L^{\min}$ be some line that has not yet been defined.*

1) If the inequality

$$\widetilde{W}\left(L^{\min} \cup (L_-(l) \cap L^{\max}) \cup \{l\}\right) \geq \widetilde{W}\left(L^{\min} \cup (L_-(l) \cap L^{\max})\right) \tag{1.29}$$

is met, then a line l can be added to the lower estimate $L^{\min}$: $L_r^{\min} = L^{\min} \cup \{l\}$, where $L_r^{\min}$ is the adjusted lower estimate.

2) If the inequality

$$\widetilde{W}\left(L^{\min} \cup (L_+(l) \cap L^{\max})\right) \geq \widetilde{W}\left(L^{\min} \cup (L_+(l) \cap L^{\max}) \cup \{l\}\right) \tag{1.30}$$

is met, then a line l can be excluded from the upper estimate $L^{\max}$: $L_r^{\max} = L_1^{\max} \setminus \{l\}$, where $L_r^{\max}$ is the adjusted upper estimate.

Proof. *Let's prove point 1. Point 2 is proved similarly. Let the specified inequality holds. It is enough to show that for any* $S \subseteq L^{\max} \setminus L^{\min}$ *the relation*

$$\widetilde{W}\left(L^{\min} \cup S \cup \{l\}\right) \geq \widetilde{W}\left(L^{\min} \cup S\right) \tag{1.31}$$

is true. If $l \in S$*, then (1.31) obviously holds. Let* $l \notin S$*. Let's split the set* S *into subsets* $S_+ = L_+(l) \cap S$ *and* $S_- = L_-(l) \cap S$*, then* $S = S_1 \cup S_2$*. Let's consider the difference of the expressions from (1.31):*

$$\widetilde{W}\left(L^{\min} \cup S \cup \{l\}\right) - \widetilde{W}\left(L^{\min} \cup S\right) = \widetilde{W}\left(L^{\min} \cup S_- \cup S_+ \cup \{l\}\right) - \widetilde{W}\left(L^{\min} \cup S_- \cup S_+\right) \geq$$
$$\geq \widetilde{W}\left(L^{\min} \cup (L_-(l) \cap L^{\max}) \cup \{l\}\right) - \widetilde{W}\left(L^{\min} \cup (L_-(l) \cap L^{\max})\right) \geq 0.$$

The inclusions $S_+ \subseteq L_+(l)$*,* $S_- \subseteq L_-(l)$*,* $S_- \subseteq L^{\max}$*, theorem 5 and the initial inequality were used here.* ■

Definition 9. *Let* $L^{\min}$ *and* $L^{\max}$ *be the current lower and upper estimates, and* (l, r) *be some pair of not yet defined lines:* $l, r \in L^{\max} \setminus L^{\min}$*,* $l \neq r$*.*

Let's define by $l \Rightarrow r$ *<u>the implication</u> «if* l *is contained in the optimal set* L^**, then* r *is also contained in* L^**», and through* $\backslash l \Rightarrow \backslash r$ *- the implication «if* l *is not contained in the optimal set* L^**, then* r *is also not contained in* L^**». Two more implications,* $l \Rightarrow \backslash r$ *and* $\backslash l \Rightarrow r$*, are defined similarly.*

The validity of the implications is established by the following theorem.

Theorem 11. *Let* $L^{\min}$ *and* $L^{\max}$ *be the current lower and upper estimates. For any pair of lines not yet defined* (l, r)*,* $l, r \in L^{\max} \setminus L^{\min}$*,* $l \neq r$*,*

1) the implication $l \Rightarrow r$ *is met if*

$$\widetilde{W}\left(L^{\min} \cup (L_-(r) \cap L^{\max}) \cup \{l, r\}\right) > \widetilde{W}\left(L^{\min} \cup (L_-(r) \cap L^{\max}) \cup \{l\}\right); \tag{1.32}$$

2) the implication $\backslash l \Rightarrow \backslash r$ *is met if*

$$\widetilde{W}\left(L^{\min} \cup (L_+(r) \cap L^{\max} \setminus \{l\})\right) > \widetilde{W}\left(L^{\min} \cup (L_+(r) \cap L^{\max} \setminus \{l\}) \cup \{r\}\right); \tag{1.33}$$

3) the implication $l \Rightarrow \backslash r$ *is met if*

$$\widetilde{W}\left(L^{\min} \cup (L_+(r) \cap L^{\max} \cup \{l\})\right) > \widetilde{W}\left(L^{\min} \cup (L_+(r) \cap L^{\max} \cup \{l, r\})\right); \tag{1.34}$$

4) the implication $\backslash l \Rightarrow r$ is met if

$$\widetilde{W}\left(L^{\min} \cup (L_{-}(r) \cap L^{\max} \setminus \{l\}) \cup \{r\}\right) > \widetilde{W}\left(L^{\min} \cup (L_{-}(r) \cap L^{\max} \setminus \{l\})\right). \qquad (1.35)$$

The proof of this statement is based on the application of theorem 5 and is similar to the proof of theorem 10.

1.6.2.1. Algorithm

Let's describe an **algorithm** for finding the optimal set of expanded lines. It uses the following variables: $L^{\min}$, $L^{\max}$ - the current lower and upper estimates of the optimal set L^*; $\overline{L}$ - the set of lines not yet defined.

Step A. Initialization of variables. Assignments are performed: $L^{\min} = \emptyset$, $L^{\max} = L$, $\overline{L} = L$.

Step B1. Inclusion and exclusion of undefined lines. Inequalities (1.29), (1.30) are checked for each undefined line $l \in \overline{L}$. If one of them holds:

1) the current estimates $L^{\min}$, $L^{\max}$ are adjusted according to theorem 10 (if both inequalities are satisfied, then the adjustment is carried out according to one of two possible options);
2) the assignment $\overline{L} = \overline{L} \setminus \{l\}$ is performed;
3) this step is repeated from the beginning.

Step B2. Checking for completion. If $L^{\min} = L^{\max}$, then the optimal set is $L^{\min}$, the algorithm terminates.

Step C. Construction of a set of implications.

1) For the current estimates $L^{\min}$, $L^{\max}$ the set of valid implications I is calculated according to theorem 11, for which, for each pair of undefined lines $(l, r) \in \overline{L} \times \overline{L}$, $l \neq r$, the implications $l \Rightarrow r$, $\backslash l \Rightarrow \backslash r$ (if l and r are mutually complementary) or $l \Rightarrow \backslash r$, $\backslash l \Rightarrow r$ (if l and r are mutually competitive) are checked. At the same time, in addition to the fair implications themselves, equivalent implications are added to the set I. For example, for $l \Rightarrow r$ such implication is $\backslash r \Rightarrow \backslash l$.
2) The set $\overline{I}$ is calculated, where $\overline{I}$ is the transitive closure of the set I, which additionally includes all the implications following from I. For example, if $I = \{1 \Rightarrow 3,\ \backslash 3 \Rightarrow \backslash 1,\ 3 \Rightarrow \backslash 5,\ 5 \Rightarrow \backslash 3\}$, then $\overline{I} = I \cup \{1 \Rightarrow \backslash 5,\ 5 \Rightarrow \backslash 1\}$.

It is worth noting that for the pair (l, r) it makes no sense to check inequalities (1.32, 1.33), if l and r are mutually competitive, and (1.34, 1.35), if l and r are mutually complementary, since they obviously do not hold (otherwise the line r would have already been defined at step B1).

Step D. Splitting a task into two subtasks.

1) For each undefined line $l \in \overline{L}$ the following values are calculated sequentially:

1.1) $F_{in}^{in}(l)$ and $F_{in}^{ex}(l)$ are sets of obviously expanded and obviously non-expanded lines when line l is extended:

$$F_{in}^{in}(l) = \left\{r \in \overline{L} \mid (l \Rightarrow r) \in \overline{I}\right\} \cup \{l\}, \quad F_{in}^{ex}(l) = \left\{r \in \overline{L} \mid (l \Rightarrow \backslash r) \in \overline{I}\right\};$$

at the same time, if $F_{in}^{in}(l) \cap F_{in}^{ex}(l) \neq \emptyset$, then a contradiction is obtained, i.e. the absence of an extension of the line l is obviously optimal;

1.2) $F_{ex}^{in}(l)$ и $F_{ex}^{ex}(l)$ - sets of obviously expanded and obviously non-expanded lines in the absence of an extension of line l:

$$F_{ex}^{in}(l) = \{r \in \overline{L} \mid (\backslash l \Rightarrow r) \in \overline{I}\}, \quad F_{ex}^{ex}(l) = \{r \in \overline{L} \mid (\backslash l \Rightarrow \backslash r) \in \overline{I}\} \cup \{l\};$$

at the same time, if $F_{ex}^{in}(l) \cap F_{ex}^{ex}(l) \neq \emptyset$, then a contradiction is obtained, i.e. the extension of line l is obviously optimal;

1.3) $F_{in}(l) = F_{in}^{in}(l) \cup F_{in}^{ex}(l)$, $F_{ex}(l) = F_{ex}^{in}(l) \cup F_{ex}^{ex}(l)$ - sets of obviously defined lines with and without extension of line l;

1.4) $c_{in}(l)$ and $c_{ex}(l)$ are bulkhead complexities for modified tasks with and without extension of line l:

$$c_{in}(l) = \begin{cases} 2^{|\overline{L} \backslash F_{in}(l)|}, & F_{in}^{in}(l) \cap F_{in}^{ex}(l) = \emptyset, \\ 0, & F_{in}^{in}(l) \cap F_{in}^{ex}(l) \neq \emptyset, \end{cases}$$

$$c_{ex}(l) = \begin{cases} 2^{|\overline{L} \backslash F_{ex}(l)|}, & F_{ex}^{in}(l) \cap F_{ex}^{ex}(l) = \emptyset, \\ 0, & F_{ex}^{in}(l) \cap F_{ex}^{ex}(l) \neq \emptyset. \end{cases}$$

2) The undefined line l^*, for which the total complexity of the two modified problems with extension and lack thereof for this line is minimal, is calculated:

$$l^* \in \operatorname*{Arg\,min}_{l \in \overline{L}} \left(c_{in}(l) + c_{ex}(l)\right).$$

3) The following two cases are considered:

3.1) line l^* is extended, in this case the set $F_{in}^{in}(l^*)$ is added to the current lower estimate, and the set $F_{in}^{ex}(l^*)$ is subtracted from the current upper estimate:

$$L^{\min} = L^{\min} \cup F_{in}^{in}(l^*),\ L^{\max} = L^{\max} \setminus F_{in}^{ex}(l^*),\ \overline{L} = \overline{L} \setminus F_{in}(l^*),$$

then, for the adjusted estimates, the algorithm is run starting from step B1; let L_{in}^* be the optimal set found;

3.2) line l^* is not extended, in this case the set $F_{ex}^{in}(l^*)$ is added to the current lower estimate, and the set $F_{ex}^{ex}(l^*)$ is subtracted from the current upper estimate:

$$L^{\min} = L^{\min} \cup F_{ex}^{in}(l^*),\ L^{\max} = L^{\max} \setminus F_{ex}^{ex}(l^*),\ \overline{L} = \overline{L} \setminus F_{ex}(l^*),$$

then, for the adjusted estimates, the algorithm is run starting from step B1; let L_{ex}^* be the optimal set found.

4) The optimal set of expanded lines is selected according to

$$L^* \in \operatorname*{Arg\ max}_{R \in \{L_{in}^*, L_{ex}^*\}} \widetilde{W}(R).$$

Substeps 3 and 4 of step D are performed as described if $F_{in}^{in}(l^*) \cap F_{in}^{ex}(l^*) = \emptyset$ and $F_{ex}^{in}(l^*) \cap F_{ex}^{ex}(l^*) = \emptyset$. If one of these ratios is incorrect, then the corresponding case cannot be optimal and should not be considered, which means that the optimal set of expandable lines coincides with the found optimal set for the only case considered.

The correctness of the algorithm follows from theorems 10-11. The benefit of using multiple implications and splitting the task into two subtasks using the selected line l^*, rather than just an arbitrary line, is especially evident in the case of the presence of a small number of large producing or consuming nodes on the market, since the inclusion/exclusion of incident lines is more likely to define some lines, as opposed to inclusion/exclusion of other lines. In general, the point of constructing implications is the ability to look «one step ahead» in order to choose the best line for dividing into subproblems at the cost of solving $O(|\overline{L}|^2)$ auxiliary problems. However, it is also possible that none of the implications will be met. Therefore, to improve the quality of the algorithm, it makes sense to initially sort the lines in descending order of their «significance» in an expert way and, in the case of an ambiguous choice, use a line with the larger value of this parameter for splitting.

1.6.2.2. Estimation of the average complexity of the algorithm

To estimate the average statistical complexity of the described algorithm, a computational experiment was conducted, similar to the experiment for the chain-type market, with piecewise linear functions of supply and demand (1.23) and quadratic functions of variable costs for increasing transmission capacity (1.24).

Let's denote by p_i^0 the equilibrium price in node $i \in N$ in the case of its isolation (with zero transmission capacity), and by Δp_i^0 the difference between the equilibrium prices for isolated nodes 0 and i: $\Delta p_i^0 = p_0^0 - p_i^0$, $i \in N \setminus \{0\}$. Let $p_{\min}^0 = \min_{i \in N} p_i^0$. Then the initial market is uniquely determined by the following parameters: N_1, N_2, $p_{\min}^0$, d_i^f $(i \in N)$, Δp_i^0, $Q_{\{i,0\}}^0$, $e_{\{i,0\}}^t$, $e_{\{i,0\}}^q$, $E_{\{i,0\}}^f$ $(i \in N \setminus \{0\})$. In this case, the equilibrium prices p_i^0, $i \in N$, are found from the system

$$\begin{cases} p_0^0 - p_i^0 = \Delta p_i^0, \ i \in N \setminus \{0\}, \\ \min_{i \in N} p_i^0 = p_{\min}^0, \end{cases}$$

and $c_i = \dfrac{d_i^f}{p_i^0}$, $i \in N$ (we assume that $p_{\min}^0 > 0$, $d_i^f > 0 \ \forall \ i \in N$).

During the experiment, the parameters $p_{\min}^0$, d_i^f, $|\Delta p_i^0|$, e_l^t, e_l^q, E_l^f characterizing the problem were randomly generated in accordance with uniform distributions (table 2), the initial transmission capacity $Q_{\{i,0\}}^0$ was taken to be 0, and for any $i \in N \setminus \{0\}$

$$\Delta p_i^0 = \begin{cases} |\Delta p_i^0|, & i \in N_1, \\ -|\Delta p_i^0|, & i \in N_2. \end{cases}$$

Table 2. Parameters of probability distributions of quantities $p_{\min}^0$, d_i^f, $|\Delta p_i^0|$, e_l^t, e_l^q, E_l^f for the star-type market

Model parameter	Minimum value	Maximum value
$p_{\min}^0$	0 (not including 0)	10
d_i^f	10	20
$\|\Delta p_i^0\|$	0	10
e_l^t	0	4
e_l^q	0	4
E_l^f	0	4

In order for the market to satisfy the FSIC, for each line $\{i, 0\} \in L$ the modified functions of marginal transport costs $e_{i0}^{nex,m}(q_{i0})$ and $e_{i0}^{ex,m}(q_{i0})$ were used: for $i \in N_1$ the type of modification is similar to (1.25, 1.26), and for $i \in N_2$ it is similar to (1.27, 1.28).

The number of nodes $|N|$ varied from 3 to 51, and the number of producing nodes coincided with the number of consuming ones. For each number of nodes 100 tasks were generated, which were solved by the described algorithm. Figure 20 shows the dependence of the number of auxiliary problems (1.6) to be solved on the number of nodes $|N|$.

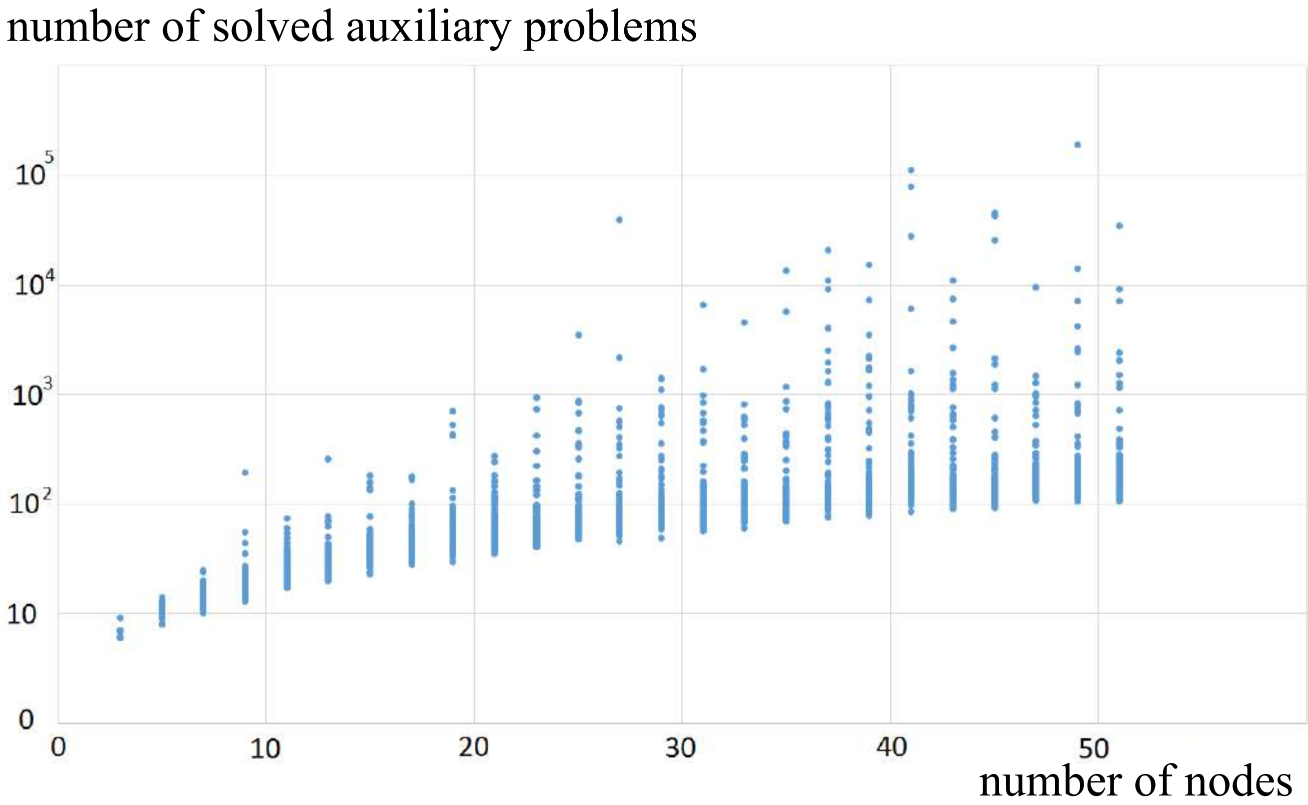

Fig. 20. Results of a numerical experiment for a star-type market. Each point corresponds to a solved problem

Let's denote the average number of auxiliary problems solved by y_{av}. The resulting dependence of y_{av} on $|N|$ is shown in figure 21. The following approximation of the average number of solved auxiliary problems was obtained by the least squares method: $\overline{y}_{av}(|N|) = 0.842\ |N|^2 - 14.776\ |N| + 54.187$ (figure 22). The corresponding coefficient of determination R^2 is 0.496.

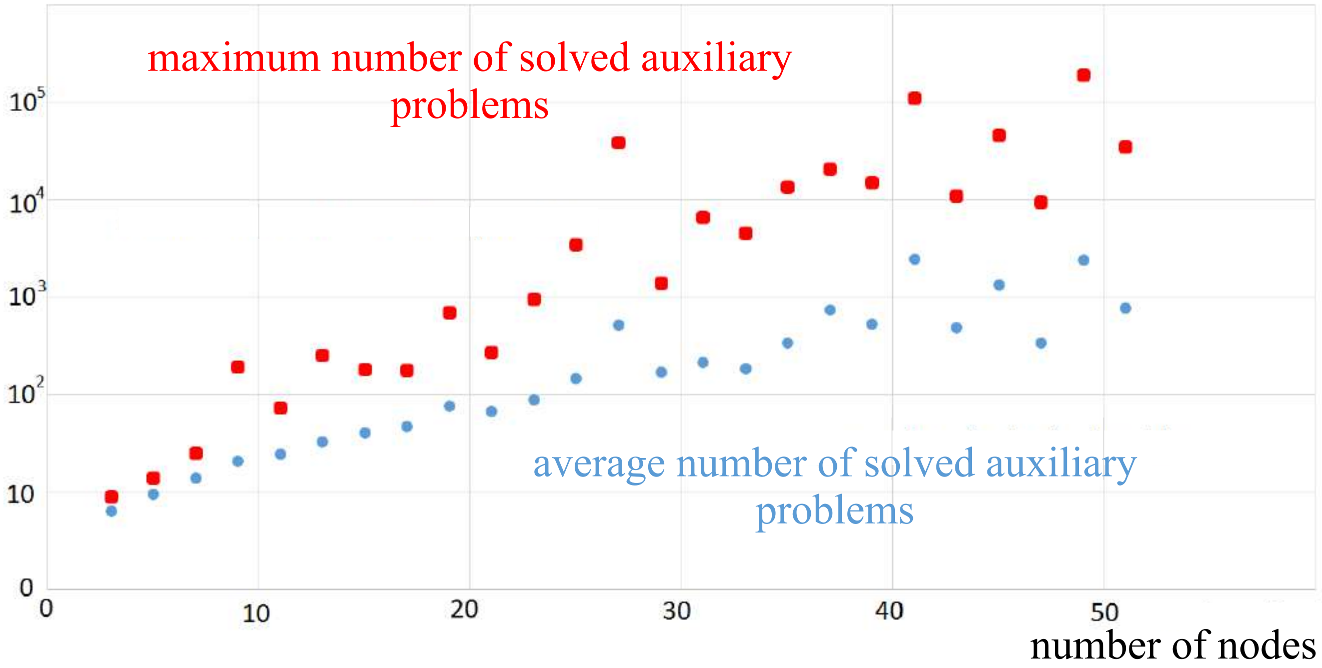

Fig. 21. Average (bottom) and maximum (top) numbers of solved auxiliary problems for the star-type market

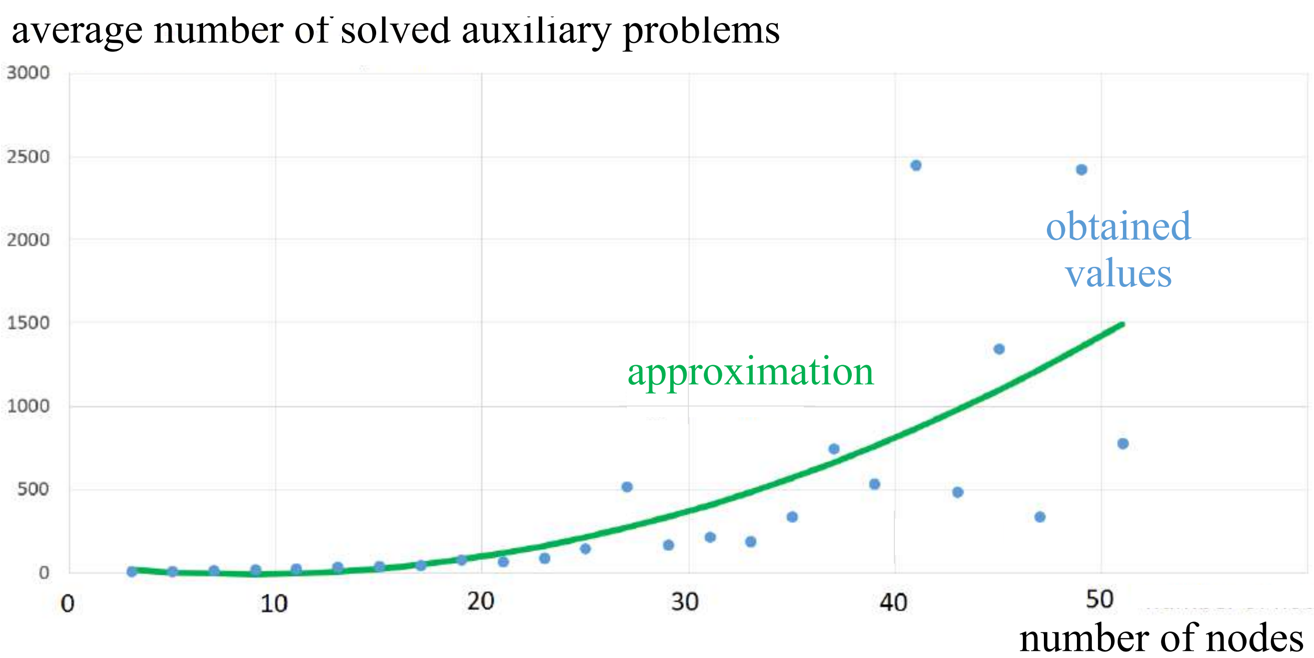

Fig. 22. Approximation of the average number of solved auxiliary problems for a star-type market

Thus, the average number of solved auxiliary problems (1.6) is no longer as well approximated by a quadratic function as for the chain-type market, which can be explained by a smaller number of generated problems (100 problems for each number of nodes instead of 1 000). However, it can be seen from the figures that the initial problem (1.5) for this type of market can also be solved in a reasonable time.

1.6.3. Star-chain-type market

Now let's consider the star-chain-type market, which is a generalization of the markets from paragraphs 1.6.1. and 1.6.2. (figure 23). This market is obtained from the star-type market by joining a chain-type market to the central node 0. Let's break down the set of lines L into subsets $L^s = \big\{\{0, i\} \in L \mid deg(i) = 1\big\}$ of the lines of the «star» part and $L^c = L \setminus L^s$ lines of the «chain» part. Here $deg(i)$ is the degree[16] of node i.

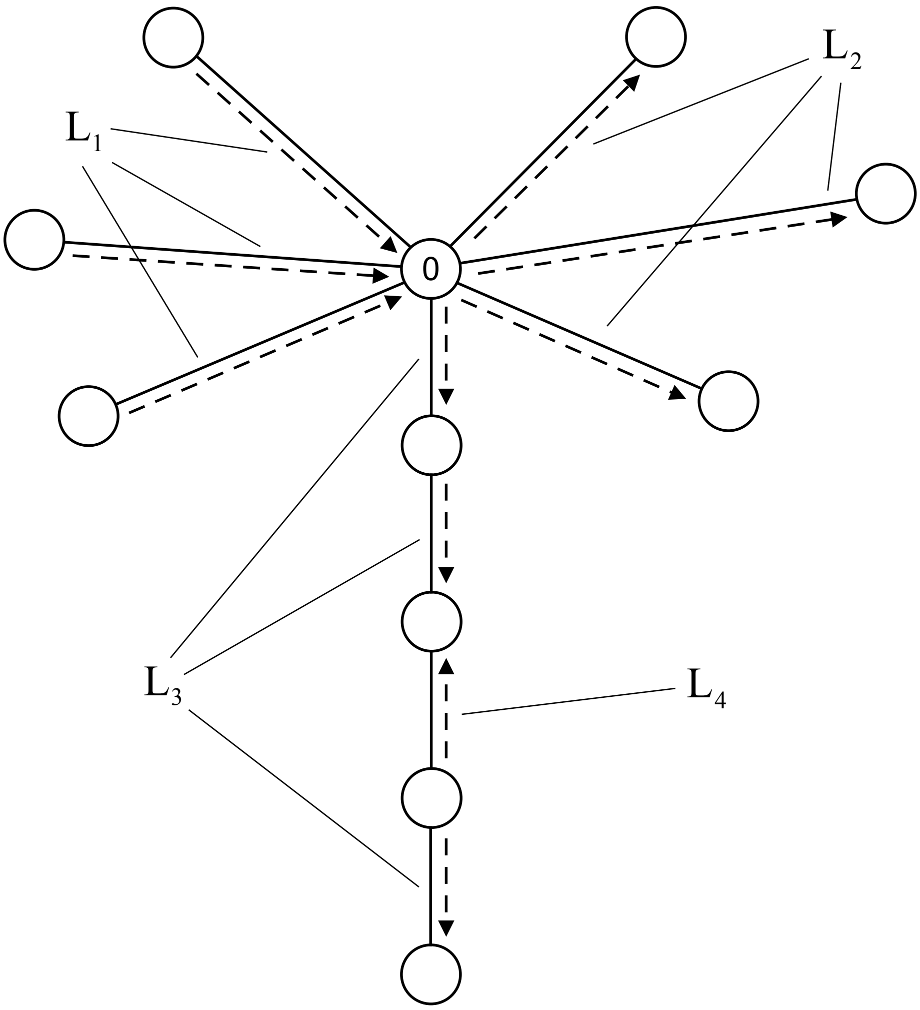

Fig. 23. An example of a star-chain-type market

Let's divide the lines of the «star» part into subsets L_1 with directions to the central node (exporting) and L_2 with directions from the central node (importing). The lines of the «chain»

[16] The degree of a node is the number of incident lines.

part, in turn, we divide into subsets L_3 with directions from the central node and L_4 with directions to the central node. The ratio of two arbitrary lines l and r, $l \neq r$, is determined according to table 3.

Table 3. The ratio of two arbitrary lines l and r for the star-chain-type market, $l \neq r$

	$r \in L_1$	$r \in L_2$	$r \in L_3$	$r \in L_4$
$l \in L_1$	competitive	complementary	complementary	competitive
$l \in L_2$	complementary	competitive	competitive	complementary
$l \in L_3$	complementary	competitive	complementary	competitive
$l \in L_4$	competitive	complementary	competitive	complementary

The algorithm for finding the optimal set of expanded lines L^* works with the lower $L^{\min}$ and upper $L^{\max}$ estimates of the set L^* and almost completely coincides with the corresponding algorithm for a star-type market. The difference lies only in one additional step, at which attempts are made to define several lines from the «chain» part. It is worth noting that theorems 10 and 11 remain true for this type of market. For any $R \subseteq L$ denote $L_k(R) = L_k \cap R$, $k \in \{1, 2, 3, 4\}$.

Theorem 12. *Let $L^{\min}$ and $L^{\max}$ be the current lower and upper estimates. Let's denote $L_k^{\min} = L_k(L^{\min})$, $L_k^{\max} = L_k(L^{\max})$, $k \in \{1, 2, 3, 4\}$. Let $S_3 \subseteq L_3^{\max} \setminus L_3^{\min}$, $S_4 \subseteq L_4^{\max} \setminus L_4^{\min}$, $S = S_3 \cup S_4$ be some sets of undefined lines from the «chain» part.*

1) If the inequality

$$\widetilde{W}\left((L_1^{\min} \cup L_3^{\min} \cup S_3) \cup (L_2^{\max} \cup L_4^{\max} \setminus S_4)\right) \geq \widetilde{W}\left((L_1^{\min} \cup L_3^{\min}) \cup (L_2^{\max} \cup L_4^{\max})\right)$$

is met and for each nonempty set $R \subset S$

$$\widetilde{W}\left((L_1^{\min} \cup L_3^{\min} \cup L_3(R)) \cup (L_2^{\max} \cup L_4^{\max} \setminus L_4(R))\right) < \\ < \widetilde{W}\left((L_1^{\min} \cup L_3^{\min}) \cup (L_2^{\max} \cup L_4^{\max})\right), \quad (1.36)$$

then the set S_3 can be added to the lower estimate $L^{\min}$, and the set S_4 can be subtracted from the upper estimate $L^{\max}$: $L_r^{\min} = L^{\min} \cup S_3$, $L_r^{\max} = L^{\max} \setminus S_4$, where $L_r^{\min}$, $L_r^{\max}$ are adjusted estimates.

2) If the inequality

$$\widetilde{W}\left((L_1^{\max} \cup L_3^{\max} \setminus S_3) \cup (L_2^{\min} \cup L_4^{\min} \cup S_4)\right) \geq \widetilde{W}\left((L_1^{\max} \cup L_3^{\max}) \cup (L_2^{\min} \cup L_4^{\min})\right)$$

is met and for each nonempty set $R \subset S$

$$\widetilde{W}\left((L_1^{\max} \cup L_3^{\max} \setminus L_3(R)) \cup (L_2^{\min} \cup L_4^{\min} \cup L_4(R))\right) < \\ < \widetilde{W}\left((L_1^{\max} \cup L_3^{\max}) \cup (L_2^{\min} \cup L_4^{\min})\right), \quad (1.37)$$

then the set S_3 can be subtracted from the upper estimate $L^{\max}$, and the set S_4 can be added to the lower estimate $L^{\min}$: $L_r^{\max} = L^{\max} \setminus S_3$, $L_r^{\min} = L^{\min} \cup S_4$, where $L_r^{\max}$, $L_r^{\min}$ are adjusted estimates.

The proof of this statement is based on the application of theorem 5 and is similar to the proof of theorem 9.

1.6.3.1. Algorithm

Let M be some maximum number of simultaneously defined lines, which is a parameter of the algorithm. The **algorithm** for finding the optimal set of expandable lines for a given type of market is obtained from the algorithm for the star-type market by adding the following step between steps B1 and B2.

Step B1.2. Inclusion and exclusion of several undefined lines from the «chain» part. We consider all possible subsets S of set $\left(L_3(L^{\max}) \setminus L_3(L^{\min})\right) \cup \left(L_4(L^{\max}) \setminus L_4(L^{\min})\right)$ for which $2 \le |S| \le M$ is satisfied, in order of ascending $|S|$: pairs of lines are considered first, then trinities, etc. For each such S:

1) the values $S_3 = L_3(S)$, $S_4 = L_4(S)$, $L_k^{\min} = L_k(L^{\min})$, $L_k^{\max} = L_k(L^{\max})$ are calculated, $k \in \{1, 2, 3, 4\}$;
2) the following inequality is checked:

$$\widetilde{W}\left((L_1^{\min} \cup L_3^{\min} \cup S_3) \cup (L_2^{\max} \cup L_4^{\max} \setminus S_4)\right) \ge \widetilde{W}\left((L_1^{\min} \cup L_3^{\min}) \cup (L_2^{\max} \cup L_4^{\max})\right),$$

   if successful, the following assignments are performed without considering the remaining subsets: $L^{\min} = L^{\min} \cup S_3$, $L^{\max} = L^{\max} \setminus S_4$, $\overline{L} = \overline{L} \setminus S$, after which there is a return to step $B1$;
3) the following inequality is checked:

$$\widetilde{W}\left((L_1^{\max} \cup L_3^{\max} \setminus S_3) \cup (L_2^{\min} \cup L_4^{\min} \cup S_4)\right) \ge \widetilde{W}\left((L_1^{\max} \cup L_3^{\max}) \cup (L_2^{\min} \cup L_4^{\min})\right),$$

if successful, the following assignments are performed without considering the remaining subsets: $L^{\max} = L^{\max} \setminus S_3$, $L^{\min} = L^{\min} \cup S_4$, $\overline{L} = \overline{L} \setminus S$, after which there is a return to step $B1$.

Note that when considering the set S in step B1.2, for any non-empty $R \subset S$ inequalities (1.36, 1.37) are met, because otherwise the lines from the set R would have already been defined (since the set R was considered before S earlier in this step), and in this case the set S would not have been considered. This means that this step is correct.

1.6.3.2. Estimation of the average complexity of the algorithm

To estimate the average statistical complexity of the described algorithm, a computational experiment was conducted, similar to the experiments in paragraphs 1.6.1. and 1.6.2., with piecewise linear functions of supply and demand (1.23) and quadratic functions of variable costs for increasing transmission capacity (1.24).

Let's denote by p_i^0 the equilibrium price in node $i \in N$ in the case of its isolation (with zero transmission capacity), and by Δp_{ij}^0 the difference between the equilibrium prices for isolated nodes i and j: $\Delta p_{ij}^0 = p_j^0 - p_i^0$, $\{i,j\} \in L$. Let $p_{\min}^0 = \min\limits_{i \in N} p_i^0$. As a result, the initial market is uniquely determined by the following parameters: $L_1 \cup L_2$, $L_3 \cup L_4$, $p_{\min}^0$, c_i $(i \in N)$, Δp_{ij}^0, $Q_{\{i,j\}}^0$, $e_{\{i,j\}}^t$, $e_{\{i,j\}}^q$, $E_{\{i,j\}}^f$ $(\{i,j\} \in L)$. In this case, the values p_i^0, $i \in N$ are found from the system

$$\begin{cases} p_j^0 - p_i^0 = \Delta p_{ij}^0, \ \{i,j\} \in L, \\ \min\limits_{i \in N} p_i^0 = p_{\min}^0, \end{cases}$$

and $d_i^f = c_i \cdot p_i^0$, $i \in N$ (we assume that $p_{\min}^0 > 0$, $c_i > 0 \ \forall \ i \in N$). In order not to be tied to one method of generating parameters, in this experiment it was decided to slightly change this method by randomly choosing coefficients c_i instead of d_i^f (this is acceptable, since the purpose of conducting computational experiments is not to compare average efficiencies for different types of markets).

Throughout the experiment, the parameters $p_{\min}^0$, c_i, $|\Delta p_{ij}^0|$, $e_{\{i,j\}}^t$, $e_{\{i,j\}}^q$, $E_{\{i,j\}}^f$ were randomly generated in accordance with a uniform distribution (table 4), the initial transmission capacity $Q_{\{i,j\}}^0$ was taken to be 0, the number of lines of the «star» part coincided with the number of lines of the «chain» part ($|L_1| + |L_2| = |L_3| + |L_4|$), in the «star» part the number of exporting lines coincided with the number of importing lines ($|L_1| = |L_2|$), for each line of the «chain» part the direction was chosen randomly with probabilities $(1/2, 1/2)$, and for any $\{i, j\} \in L$ the sign Δp_{ij}^0 corresponded to the selected directions:

$$\Delta p_{ij}^0 = \begin{cases} |\Delta p_{ij}^0|, & \text{the direction from node } i \text{ to node } j, \\ -|\Delta p_{ij}^0|, & \text{the direction from node } j \text{ to node } i. \end{cases}$$

Table 4. Parameters of probability distributions of quantities $p_{\min}^0$, c_i, $|\Delta p_{ij}^0|$, $e_{\{i,j\}}^t$, $e_{\{i,j\}}^q$, $E_{\{i,j\}}^f$ for the star-chain-type market

Model parameter	**Minimum value**	**Maximum value**
$p_{\min}^0$	0 (not including 0)	10
c_i	1	5
$\|\Delta p_{ij}^0\|$	0	10
$e_{\{i,j\}}^t$	0	4
$e_{\{i,j\}}^q$	0	4
$E_{\{i,j\}}^f$	0	4

In order for the market to satisfy the FSIC, in the experiment, for each line $\{i, j\} \in L$ modified functions of marginal transmission costs $e_{ij}^{nex,m}(q_{ij})$ and $e_{ij}^{ex,m}(q_{ij})$ were used: in the case of the selected direction from node i to node j the type of modification is similar (1.25, 1.26), otherwise it is similar to (1.27, 1.28).

The number of nodes $|N|$ varied from 5 to 73. For each $|N|$ 10 000 problems were generated, which were solved by the described algorithm with $M = 10$. Let's denote the average number of auxiliary problems solved by y_{av}. The resulting dependence of y_{av} on $|N|$ is shown in figure 24. The following approximation of the average number of solved auxiliary problems was obtained by the least squares method: $\overline{y}_{av}(|N|) = 0.0175\,|N|^2 + 3.4262\,|N| - 3.3184$ (figure 25). The corresponding coefficient of determination R^2 is equal to 0.9998.

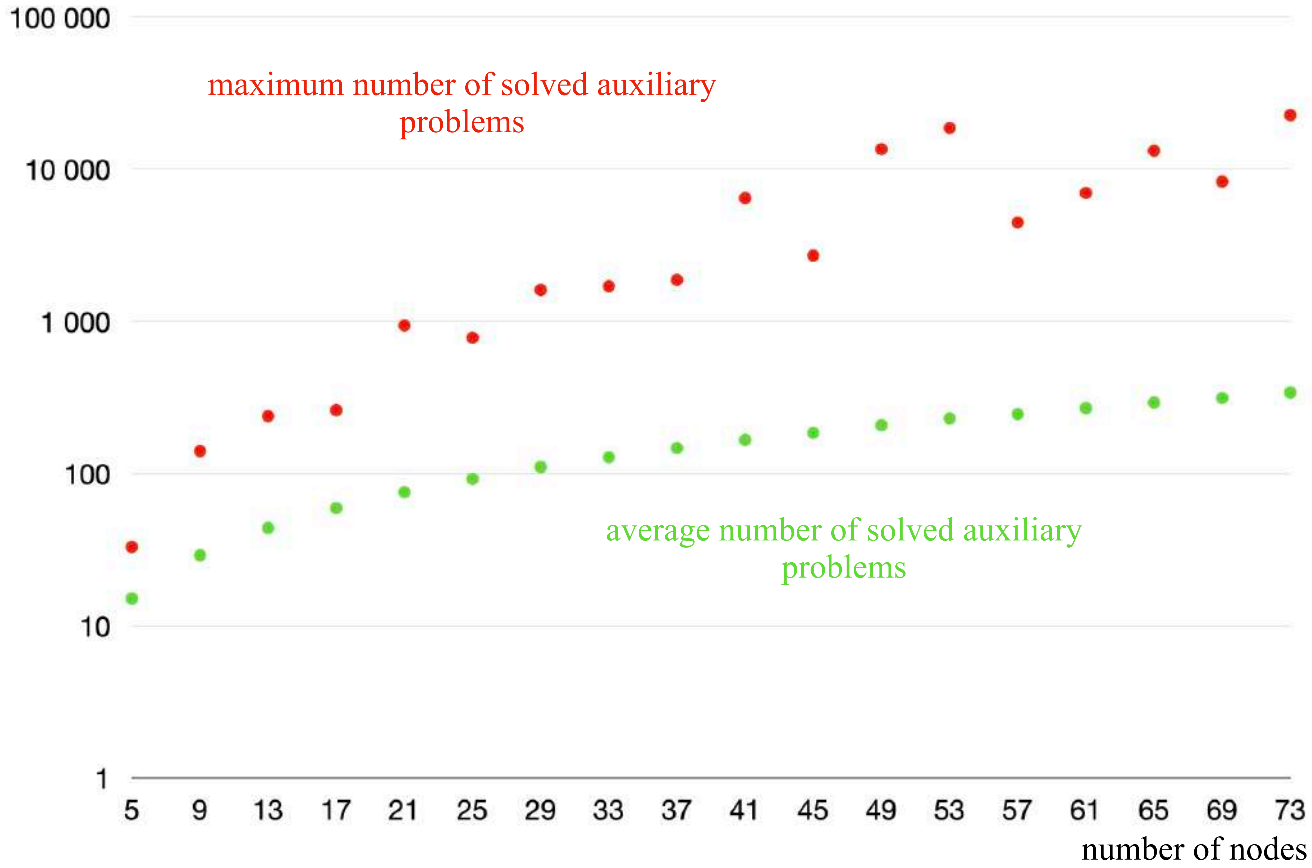

Fig. 24. Average (bottom) and maximum (top) numbers of solved auxiliary problems for the star-chain-type market

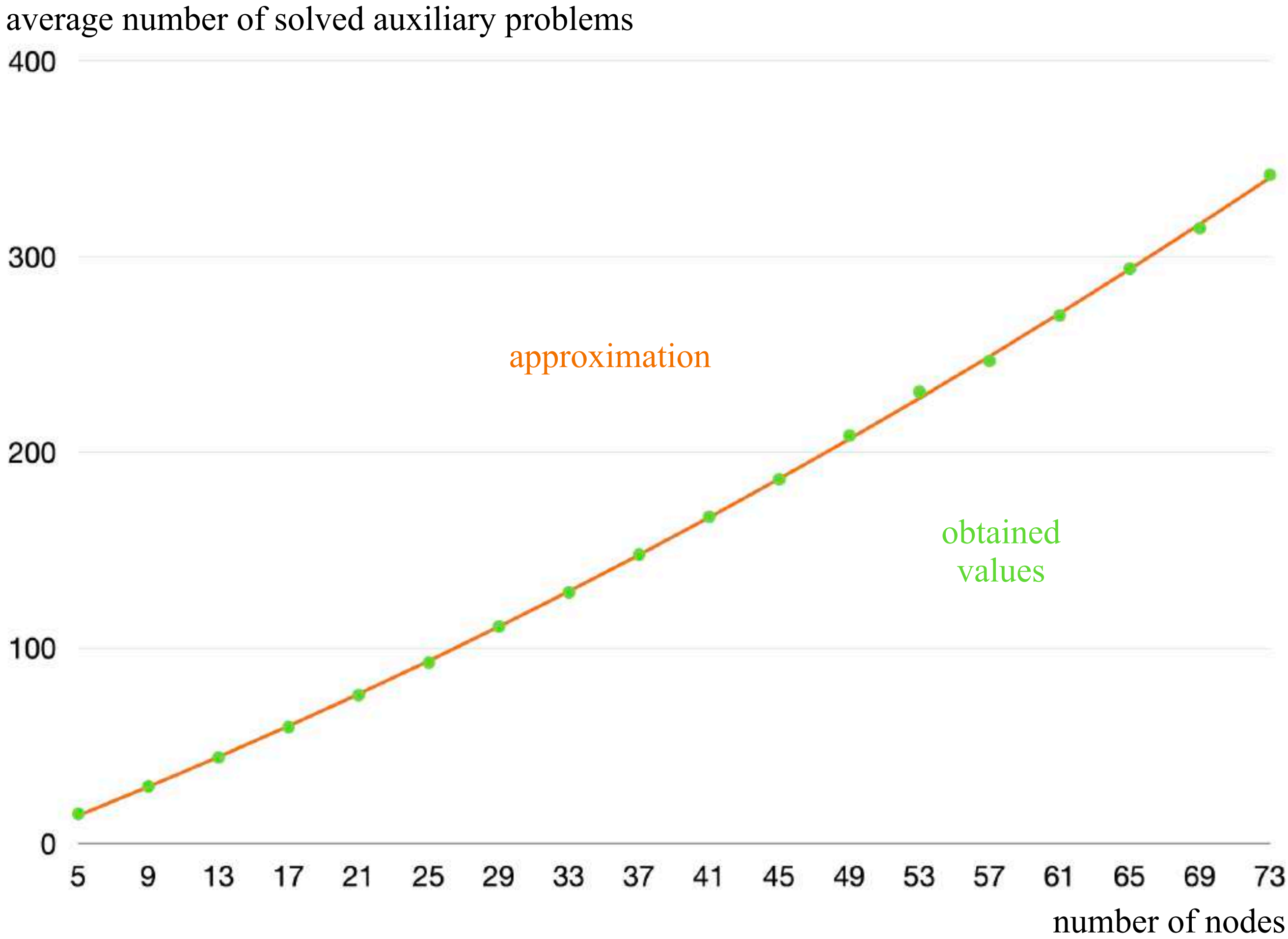

Fig. 25. Approximation of the average number of solved auxiliary problems for a star-chain-type market

It can be seen from the figures that the average number of solved auxiliary problems (1.6) is very well approximated by the resulting quadratic function. Such high accuracy compared to the approximations from paragraphs 1.6.1. and 1.6.2. can be explained, firstly, by the larger number of generated problems (10 000 problems for each number of nodes instead of 1 000 and 100 respectively), вand secondly, by a slightly modified experiment.

1.6.4. Tree-type market

In the master's thesis by Silaev I. I. [56] the market of the general tree-type is considered, for which an algorithm for finding the optimal set of expanded lines L^* is proposed[17]. This algorithm is a generalization of the algorithm for the star-type market (see paragraph 1.6.2.). Theorems 10

[17] The results described in this paragraph do not belong to the author of this work and are provided for completeness.

and 11 remain true for the tree-type market.

Let's divide the set of all lines L into the minimum number of disjoint subsets $L_1, ..., L_K$ in such a way that

1) each subset L_k, together with incident nodes, formed a chain-type market (i.e., a connected market in which the degree of each node does not exceed 2);
2) for any two different subsets L_{k_1} and L_{k_2}, $k_1 \neq k_2$, the corresponding sub-markets could only intersect with the end nodes.

Figure 26 shows an example of such a split. Let's divide each set L_k, $k \in \{1, ..., K\}$, into two subsets L_k^1, L_k^2 in accordance with the directions of the equilibrium flows (the directions of the equilibrium flows coincide if the lines belong to the same subset, and are opposite otherwise). For any $m \in \{1, 2\}$ and $k \in \{1, ..., K\}$ the set L_k^m has the following property: $\forall\ l, r \in L_k^m$, $l \neq r$, the equalities $L_+(l) \cup \{l\} = L_+(r) \cup \{r\}$, $L_-(l) = L_-(r)$ are valid. Let's denote by $L_-(L_k^1)$, $L_-(L_k^2)$ the sets of competitive lines for lines from the sets L_k^1 and L_k^2 respectively. For any $k \in \{1, ..., K\}$, $R \subseteq L$ let's denote $L_k^1(R) = L_k^1 \cap R$, $L_k^2(R) = L_k^2 \cap R$.

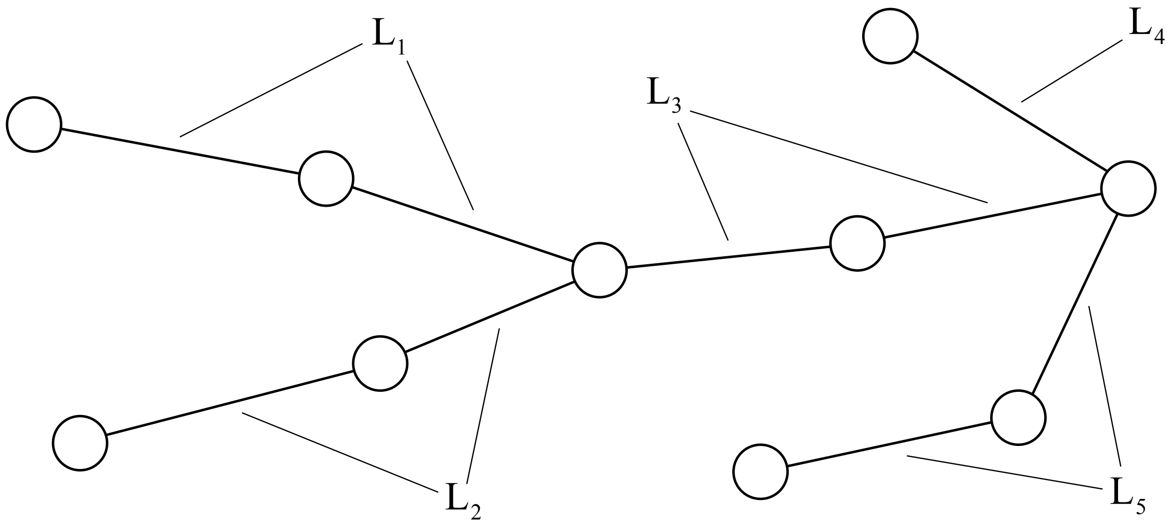

Fig. 26. An example of a tree-type market

Unlike the star-chain-type market, the tree-type market can contain several «chain» parts, for each of which an analogue of theorem 12 can be applied, allowing several lines to be determined simultaneously when the corresponding inequalities hold. Here is the corresponding statement.

Theorem 13. *Let $L^{\min}$ and $L^{\max}$ be the current lower and upper estimates, and $k \in \{1, ..., K\}$. Let $S_k^1 \subseteq (L^{\max} \setminus L^{\min}) \cap L_k^1$, $S_k^2 \subseteq (L^{\max} \setminus L^{\min}) \cap L_k^2$, $S_k = S_k^1 \cup S_k^2$ be some sets of undefined lines included in L_k.*

1) If the inequality

$$\widetilde{W}\left((L^{\min} \cup S_k^1) \cup (L_-(L_k^1) \cap L^{\max} \setminus S_k^2)\right) \geq \widetilde{W}\left(L^{\min} \cup (L_-(L_k^1) \cap L^{\max})\right) \qquad (1.38)$$

is met and for each nonempty set $R \subset S_k$

$$\widetilde{W}\left((L^{\min} \cup L_k^1(R)) \cup (L_-(L_k^1) \cap L^{\max} \setminus L_k^2(R))\right) <$$
$$< \widetilde{W}\left(L^{\min} \cup (L_-(L_k^1) \cap L^{\max})\right),$$

then the set S_k^1 can be added to the lower estimate $L^{\min}$, and the set S_k^2 can be subtracted from the upper estimate $L^{\max}$: $L_r^{\min} = L^{\min} \cup S_k^1$, $L_r^{\max} = L^{\max} \setminus S_k^2$, where $L_r^{\min}$, $L_r^{\max}$ are adjusted estimates.

2) If the inequality

$$\widetilde{W}\left((L^{\min} \cup S_k^2) \cup (L_-(L_k^2) \cap L^{\max} \setminus S_k^1)\right) \geq \widetilde{W}\left(L^{\min} \cup (L_-(L_k^2) \cap L^{\max})\right) \qquad (1.39)$$

is met and for each nonempty set $R \subset S_k$

$$\widetilde{W}\left((L^{\min} \cup L_k^2(R)) \cup (L_-(L_k^2) \cap L^{\max} \setminus L_k^1(R))\right) <$$
$$< \widetilde{W}\left(L^{\min} \cup (L_-(L_k^2) \cap L^{\max})\right),$$

then the set S_k^1 can be subtracted from the upper estimate $L^{\max}$, and the set S_k^2 can be added to the lower estimate $L^{\min}$: $L_r^{\max} = L^{\max} \setminus S_k^1$, $L_r^{\min} = L^{\min} \cup S_k^2$, where $L_r^{\max}$, $L_r^{\min}$ are adjusted estimates.

Let M be some maximum number of simultaneously defined lines, which is a parameter of the algorithm. The **algorithm** for finding the optimal set of expandable lines for a given type of market is obtained from the algorithm for the star-type market by adding the following step between steps B1 and B2.

Step B1.2. Inclusion and exclusion of several undefined lines from one «chain»-part. All possible sets of undefined lines $S \subseteq L^{\max} \setminus L^{\min}$, such that $2 \leq |S| \leq M$ and $\exists$ $k \in \{1, ..., K\}$: $S \subseteq L_k$, are considered in increasing $|S|$ order: pairs of lines are considered first, then triples, etc. For each such S and the corresponding k:

1) the values $S_k^1 = L_k^1(S)$, $S_k^2 = L_k^2(S)$ are calculated;
2) the inequality (1.38) is checked, if successful, without considering the remaining subsets,

assignments $L^{\min} = L^{\min} \cup S_k^1$, $L^{\max} = L^{\max} \setminus S_k^2$, $\overline{L} = \overline{L} \setminus S$ are performed, after which a return to step $B1$ occurs;

3) the inequality (1.39) is checked, if successful, without considering the remaining subsets, the assignments $L^{\max} = L^{\max} \setminus S_k^1$, $L^{\min} = L^{\min} \cup S_k^2$, $\overline{L} = \overline{L} \setminus S$ are performed, after which a return to step $B1$ occurs.

The correctness of this step is explained in the same way as for the star-chain-type market. To estimate the average statistical complexity of the described algorithm, a computational experiment was conducted similar to the experiment in paragraph 1.6.3. 1 000 problems were generated for each number of nodes. As a result of the experiment, the following linear approximation of the average number of solved auxiliary tasks was obtained depending on the number of nodes in the market: $\overline{y}_{av}(|N|) = 3.6127\ |N| - 4.3794$. The corresponding coefficient of determination R^2 is 0.9999.

1.6.5. Comparison of algorithms for different types of markets

If we compare the considered types of markets in terms of the capabilities of algorithms, it is important to note the following feature of markets that contain parts of the «chain» part. For each «chain», the set of lines belonging to it is divided into 2 classes in accordance with the directions of equilibrium flows, and any line outside this «chain» is complementary for all lines of one of these classes and competitive for all lines of another class. It is these properties that explain the existence of theorems 9, 12 and 13 for such markets, which allow several lines to be defined simultaneously. Note that there is no analog of these theorems for the star-type market. Thus, when using the developed algorithms to optimize markets with «chain» parts, there is an advantage in the form of an additional possibility of defining several lines using the stated theorems. However, of course, this does not mean at all that the number of auxiliary problems solved for such markets is always less, since ultimately the number of auxiliary problems solved (1.6) is strongly influenced by the characteristics of nodes and lines, not only by the transport structure of the market.

Conclusions to the first chapter

In this chapter, the problem of optimizing the transport system of a multi-node energy market of a homogeneous product in terms of maximizing social welfare was investigated, and its NP-hardness was proved. It is shown that the initial problem is reduced to the problem of finding the optimal set of expandable lines, to which an auxiliary problem with a fixed set of expandable lines acts as a sub-problem.

The concept of competitive equilibrium is introduced. It is shown how it is related to the solution of an auxiliary problem. To solve the latter, an algorithm has been developed, the complexity of which for the case of piecewise linear initial functions depends quadratically on the number of nodes in the market.

Two particular cases of a chain-type market are considered: with zero initial capacity and with monotonous initial equilibrium prices. Polynomial algorithms for solving the initial problem have been developed for them. They have quadratic and cubic difficulties, respectively.

The markets of the «chain», «star», «star-chain» types are considered separately if the flow structure invariance condition is met. Algorithms for solving the initial problem have been developed for them. These algorithms are based on the use of the properties of the social welfare function, which enable finding lines the expansion of which (or, conversely, lack of expansion) is obviously optimal. Computational experiments were conducted to estimate the average complexity of these algorithms. They showed that for each type of market considered, the average number of auxiliary problems solved is well approximated by a quadratic function depending on the number of nodes in the market. The algorithm and the results of similar computational experiments for the tree-type market, obtained in another study, are also presented.

Based on the results of the experiments, it can be concluded that the developed algorithms allow solving the initial problem (1.5) in an acceptable time even for markets with a large number of nodes, which was not at all obvious at the initial stage of the study of this problem, since it is NP-hard in the general case.

Chapter 2. Application of the developed algorithms to assess the prospects of gasification of Russian regions

This chapter discusses issues related to the practical use of the model and algorithms described in the first chapter. In order for the proposed model to be used for planning the development of transport systems in real energy markets, it is necessary to be able to assess the initial parameters of the model for such markets - the functions of production costs, consumption utility and transmission costs. It is also necessary to be able to identify promising transport lines and take into account the possibility of their construction in the model.

In general, the way to solve the above-mentioned problems depends on the type of energy resource used and the specifics of the market. This chapter examines the Russian natural gas market, but the proposed models and methods can be adapted for other markets.

Paragraph 2.1. describes the main areas of application of natural gas, paragraph 2.2. provides the structure of natural gas consumption in the Russian Federation.

Paragraph 2.3. is devoted to the problem of estimating the transport cost function for a new gas pipeline. Two types of gas pipelines are considered: main and distribution pipelines. For each type, an estimate of the transmission cost function is derived, depending on the capacity and length of the gas pipeline.

In paragraph 2.4. the problem of modelling the production cost function for a gas field is considered. An overview of one well-known model describing the dynamics of natural gas production from a gas deposit, depending on the rate of commissioning of producing wells and the selected technological mode of their operation, is given. It is shown how, based on these parameters, the total costs for the operation of a gas deposit are calculated. A simplified model for estimating the production cost function is also proposed, and an appropriate estimate is derived.

Paragraph 2.5. is devoted to the problem of forecasting the demand for natural gas in a node that is an arbitrary non-gasified region or municipality of the Russian Federation. The main potential gas consumers are divided into several groups. A mathematical model is described for each group to evaluate the corresponding component of the demand function[18].

In the last paragraph 2.6. the developed algorithms and methods are used to assess the prospects of gasification of Irkutsk Oblast. The possibility of connecting thermal power plants

[18] Although the initial parameters of the model describing consumers are not demand functions, but consumption utility functions, there is a one-to-one correspondence for these functions (1.11).

and boiler houses in the region to the main gas pipeline «Power of Siberia» is being considered, while determining the optimal plan for the development of the gas network in terms of maximizing social welfare. For calculations, real data from official sources are used: a database of indicators of municipalities published on the website of the Federal State Statistics Service [73], a database of the socio-economic situation of the subjects of the Russian Federation [74], as well as decrees of the governor of Irkutsk Oblast on approval of schemes and programmes for the development of the electric power industry [2–4].

2.1. About natural gas

Natural gas is one of the most economically significant energy resources in Russia and on the planet. The advantage of natural gas over other types of organic fuels lies in its low cost and environmental friendliness, and its proven reserves in Russia, according to the statement of Energy Minister Alexander Novak, will last for more than 100 years [65]. At the same time, Russia is the world leader in natural gas reserves (in 2019, its share was 19%, [66]).

The main use of natural gas is related to energy [75]. It can be used directly as an energy source or converted into other types of energy. The main applications of natural gas as a direct source of energy are industrial production, transportation and household consumption. As for conversion to other types of energy, in Russia natural gas is the main fuel for thermal power plants, combined heat and power plants and boiler houses. The use of alternative fuels (mainly coal) is generally associated with the lack of access to pipeline gas in specific areas of the country.

Natural gas is also widely used as a raw material for processing into other fuels and in the production of chemical and other non-fuel products. In addition to methane, which is the main component of natural gas (80-90%) used as fuel, it contains valuable impurities: pentane, hexane, butane, propane, ethane, nitrogen and helium. These elements are separated at specially created gas processing plants, one of the largest of which in Russia is the Amur Gas Processing Plant, launched in 2021 [76]. The possibilities of using the obtained substances are extensive: from the production of plastics and paints to the creation of parts for MRI machines and liquid crystal screens.

It is worth noting that the use of methane instead of gasoline as an automobile fuel is becoming increasingly popular due to its low price (in terms of the cost of fuel per kilometer, methane is 2-3 times cheaper), which indicates an increase in the potential of natural gas. The average annual growth in global natural gas consumption in 2008-2019 was about 2.5% [66].

2.2. Natural gas consumption in the Russian Federation

According to official data [75], in 2019[19], the volume of consumption of Russian natural gas amounted to 851.9 million tons, of which 254.6 million tons were exported, the rest was domestic consumption. The structure of domestic consumption of natural gas is shown in figure 27.

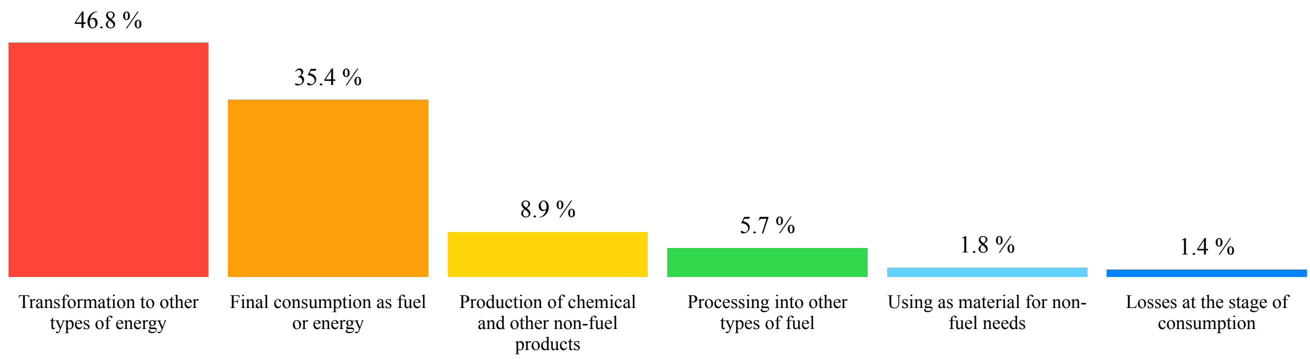

Fig. 27. Structure of natural gas consumption by industrial production in the Russian Federation in 2019

The figure shows that over 90% of natural gas is spent on conversion to other types of energy (primarily electrical and thermal), final consumption as fuel or energy, and the production of chemical and other non-fuel products. Also, about 6% of the gas consumed is processed into other fuels and about 2% is used as a material for non-fuel needs. Gas losses are less than 1.5%.

Of the total final consumption of natural gas as fuel or energy, the following main components can be distinguished (figure 28): supply to the population (43.2%), industrial production (31%), transportation and storage (18.8%), construction (2.5%).

In industrial production, 71.8% of natural gas is consumed by processing industries, 27.3% - in the extraction of minerals (mainly crude oil, natural gas and metal ores), less than 1 % - by other industries (figure 29).

[19] At the time of the study, data for later years were not available.

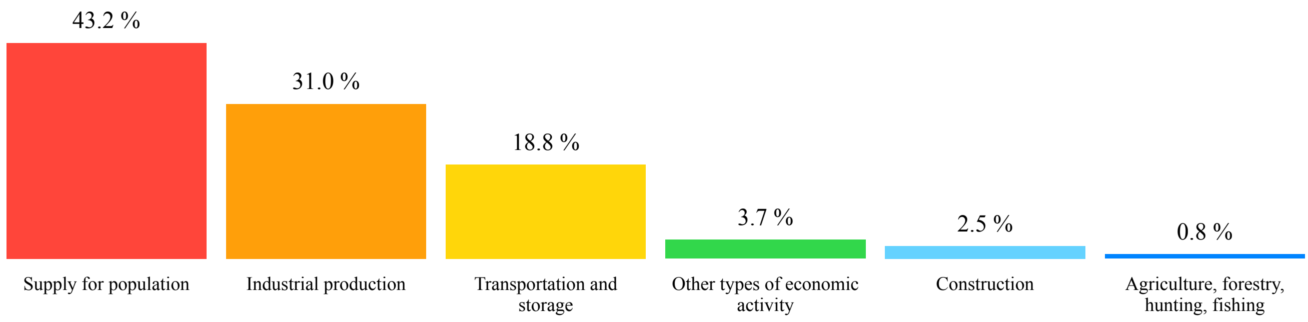

Fig. 28. The structure of final consumption of natural gas as fuel or energy, divided by type of economic activity in the Russian Federation in 2019

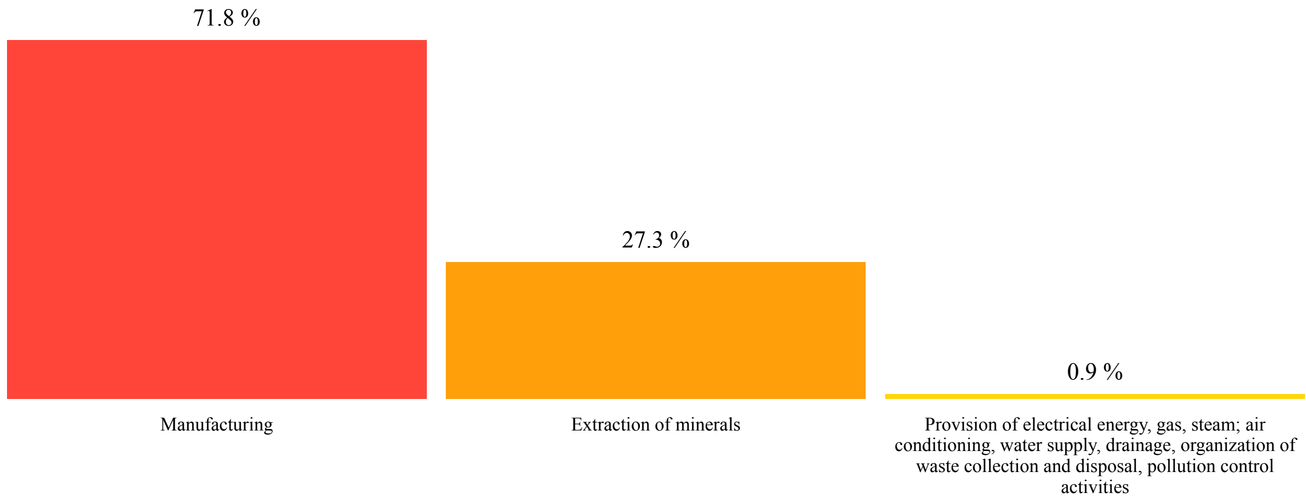

Fig. 29. The structure of natural gas consumption by industrial production in the Russian Federation in 2019

2.3. Estimation of the transmission cost function for a new gas pipeline

Later in this chapter, for convenience, the volume of natural gas and energy are measured in tons of conventional fuel (t.c.f.). Conventional fuel is the standard unit of accounting for the calorific value of fuel. It is assumed that when 1 conventional fuel is completely burned, 29.3 MJ (7 000 kcal) of heat is released. To convert the volume of natural gas into an equivalent mass of conventional fuel, the coefficient $k_{g\to c.f.} = 1.154 \times 10^{-3}$ t.c.f./m^3 is used, and to convert the mass of conventional fuel into energy, the coefficient $k_{c.f.\to j} = 29.3 \times 10^9$ (J/t.c.f.).

2.3.1. Main gas pipelines

The main gas pipeline is used to transport large volumes of natural gas over long distances. It consists of large diameter pipes in which high pressure is maintained due to compressor stations located on the path of the gas pipeline. According to the set of rules [1] approved by the Ministry of Construction and Housing and Communal Services of the Russian Federation and establishing the basic requirements for the production and acceptance of construction and installation works during the construction and reconstruction of the linear part of the main pipelines, the main costs during the construction of the main gas pipeline can be divided into 5 categories:

- costs for the purchase and delivery of pipes;
- costs of off-piste work (construction of access roads, power lines, energy supply systems and residential towns of builders; preparation of construction sites and equipment storage sites; construction of bridges along the route for delivery of construction equipment and cargo; other work);
- the cost of carrying out long-distance along-track work, minus the cost of purchasing and delivering pipes (clearing the construction strip from forest vegetation, draining the construction strip, arranging crossings over water barriers, building crossings under roads and railways, cutting steep longitudinal slopes, construction of underwater and tunnel crossings, protection of the territory of the construction strip from adverse natural phenomena, earthworks, pipe laying and welding, quality control, etc.);
- costs for the construction of compressor stations;
- other costs (for design, lease or purchase of land, etc.).

A significant proportion of all costs are fixed costs that do not depend on the total capacity of the gas pipeline. Such costs include the costs of designing, renting or buying land, most of the costs of off-piste and along-track work, as well as fixed costs for the construction of compressor stations. We assume that the remaining costs are proportional to the capacity, and also that all these costs are proportional to the length of the gas pipeline. The cost of operating a gas pipeline can be considered proportional to its capacity and length.

When estimating the function of transmission costs, we neglect the influence of such factors as the geographical location of the territory of the gas pipeline, its relief, the average annual temperature, the height difference, the size of the water sections crossed by the route, etc. In this case, the transmission cost function of the main gas pipeline depends only on its capacity and length and has the following form:

$$\overline{E}^m(q,l) = \begin{cases} 0, & |q| = 0, \\ (e^{m,f} + e^{m,l} \cdot |q|) \cdot l, & |q| > 0. \end{cases} \quad (2.1)$$

Here $|q|$ is the volume of natural gas transferred (t.c.f./year), l is the length of the gas pipeline (km). The value of $\overline{E}^m(q,l)$ is an estimate of total transmission costs (rubles/year), which takes into account the cost of natural gas transmission and the reduced capital costs for the construction of the gas pipeline. Capital expenditures are adjusted to an annual period of time, taking into account the expected service life of the gas pipeline, the interest rate on bank deposits and the inflation rate (see paragraph 1.1.2. and (1.1)). It is assumed that the capacity of the gas pipeline is proportional to the transmitted volume $|q|$ and exceeds it, creating a reserve that takes into account the annual uneven consumption.

The parameters $e^{m,f}$ and $e^{m,l}$ should be set expertly, since the publicly available data related to the cost of construction of individual main gas pipelines of the Russian Federation are very general and vary greatly for different projects. For this study, open data on the «Power of Siberia» gas pipeline are used [77,78]. As of April 27, 2018, the project cost amounted to 1.1 trillion rubles, the planned transmission volume is 42[20] billion m^3 (48.47 million t.c.f.) per year, the length of the gas pipeline is 2963.5 km. Thus, the cost of building a kilometer of the gas pipeline amounted to 371 million rubles on the specified date. Assuming the share of fixed costs of 25%, the service life of the gas pipeline of 50 years and the equality of expected interest rates on bank deposits to inflation levels, we obtain that in prices at the beginning of 2022, the annual cost of building a kilometer of the gas pipeline consists of a fixed component equal to 2.25 million rubles and a variable component equal to 0.139 rub. for each ton of conventional fuel in the capacity.

To estimate the cost of gas transmission, we will take into account the statement of a representative of Gazprom [79], according to which the cost of pumping 1 thousand m^3 of natural gas from independent producers per 100 km in 2017 was 69.2 rubles. Thus, in prices at the beginning of 2022, the transportation of 1 t.c.f. of natural gas per 1 km costs 0.739 rubles.

As a result, the following parameter estimates are obtained: $e^{m,f} = 2.25 \times 10^6$ (rub./(km×year)), $e^{m,l} = 0.878$ (rub./(t.c.f.×km)), while the function of transmission costs for

[20] Taking into account the gas consumption of the Amur gas processing plant.

the main gas pipeline in prices at the beginning of 2022 takes the form

$$\overline{E}^{m}(q,l)=\begin{cases}0, & |q|=0,\\ (2.25\times 10^{6}+0.878\cdot |q|)\cdot l, & |q|>0.\end{cases} \tag{2.2}$$

2.3.2. Gas pipelines of distribution networks

Gas pipelines of distribution networks (hereinafter referred to as distribution pipelines) are designed to transport natural gas from gas distribution stations to the final consumer. Such gas pipelines have a short length and small diameter, and the pressure in gas distribution stations is reduced to the level necessary to supply consumers. Compared to main pipelines, distribution gas pipelines are less efficient, and the requirements for their construction are much lower, so the construction and operation costs for these two types of gas pipelines differ greatly.

Estimation of the transmission cost function for a new gas pipeline in a form similar to the assessment for the main gas pipeline:

$$\overline{E}^{d,1}(q,l)=\begin{cases}0, & |q|=0,\\ (e^{d,f}+e^{d,l}\cdot |q|)\cdot l, & |q|>0.\end{cases} \tag{2.3}$$

The parameters $e^{d,f}$ and $e^{d,l}$ should be set expertly. To assess the parameters $e^{d,f}$ and $e^{d,l}$, we will take into account data from the official website of the unified gasification operator of the Russian Federation [80], according to which, in accordance with the connection programme «Social Gasification», the approximate cost of construction and installation works for underground laying of a kilometer of a distribution gas pipeline with a capacity limit of up to 7 m^3/hour in Irkutsk Oblast is 1.134[21] million rubles. Data on other areas are also available. Assuming a fixed cost share of 25%[22], the service life of the distribution gas pipeline of 30 years, the equality of expected interest rates on bank deposits to inflation levels and disregard for the costs of operating the gas pipeline, we obtain that in prices at the beginning of 2022, the estimate of the transmission cost

[21] For May 2022.

[22] The value of the fixed cost share of 25% is highlighted only as an example, this value should be set expertly. The estimate of the transport cost function obtained below based on this value is not used in analyzing the prospects for gasification of Irkutsk Oblast.

function for the distribution gas pipeline in Irkutsk Oblast has the following form:

$$\overline{E}^{d,1}(q,l) = \begin{cases} 0, & |q| = 0, \\ (8.595 \times 10^3 + 364.4 \cdot |q|) \cdot l, & |q| > 0. \end{cases}$$

Thus, $e^{d,f} = 8.595 \times 10^3$ (rub./(km×year)), $e^{d,l} = 364.4$ (rub./(t.c.f.×km)).

We will also construct an estimate of the transport cost function for a distribution gas pipeline without a fixed component:

$$\overline{E}^{d,2}(q,l) = e^d \cdot |q| \cdot l. \tag{2.4}$$

This estimate can later be used instead of the estimate $\overline{E}^{d,1}(q,l)$ in cases where, according to the specifics of the model, the transport cost function should be convex (functions $\overline{E}^{m}(q,l)$ and $\overline{E}^{d,1}(q,l)$ are not convex in q due to the presence of fixed components $e^{m,f}$ and $e^{d,f}$ respectively). It is implied that the accuracy of this estimate is worse[23] than the estimates $\overline{E}^{d,1}(q,l)$. For Irkutsk Oblast, it takes the following form:

$$\overline{E}^{d,2}(q,l) = 485.8 \cdot |q| \cdot l,$$

that is, $e^d = 485.8$ (rubles./(t.c.f.×km)).

It is worth noting that all the estimates of the transmission cost function obtained are in some way overestimated, since they include the profits of contractors engaged in the construction of the gas pipeline. It also makes sense to take this profit into account in the function of social welfare, for which it is recommended to set the evaluation parameters expertly. For the purposes of this study, the estimates obtained are considered acceptable.

2.4. Estimation of the production cost function for a gas deposit

The development of a natural gas field is a complex technological process that includes the following main stages: exploration, assessment of gas reserves, preparation for industrial production, industrial production, liquidation. Carrying out each of these stages requires significant economic investments, the amount of which strongly depends on the characteristics of a particular

[23] When making estimates correctly, during which the range of variation of the variable q is taken into account, the error of the estimate $\overline{E}^{d,2}(q,l)$ relative to the estimate $\overline{E}^{d,1}(q,l)$ does not exceed the value $e^{d,f} \cdot l$.

field.

The longest and most expensive are the stages of preparation for industrial production and industrial production itself, during which wells are built and natural gas reserves are extracted from the bowels of the earth. The physical basis of the principle of gas extraction is the difference (depression) between the pressures in the gas-bearing reservoir and at the bottom of the producing well, causing the movement of gas from the reservoir to the bottom of the well and its further elevation to the surface of the earth.

It is worth noting that the process of field development is not easy (Yushkov I. R., Khizhniak G. P., Ilyushin P. Yu., 2013, [57]; Minkhanov I. F., Dolgikh S. A., Varfolomeev M. A., 2019, [58]). So, in addition to producing wells that perform the main function of gas extraction, there are injection, exploration, observation, control, evaluation and reserve wells, while the shape and inclination of the wells are divided into vertical, directional, horizontal, multi-barrel and multi-hole. The systems for placing wells on the gas-bearing area also differ: square, triangular, circular, etc.

Among the modes of operation of gas-bearing formations, two main ones should be noted: gas and water-pressure [57]. In the gas regime, the flow of gas to the faces of producing wells occurs only due to its elastic expansion, therefore the volume of the pore space occupied by gas practically does not change. Under the water pressure regime, the flow of gas to the well faces is additionally influenced by groundwater surrounding the deposit and displacing gas, thereby slowing down the rate of reduction of reservoir pressure and maintaining the flow rate (gas production rate) at a high level.

Also, when developing a gas field, various methods of increasing the flow rate can be used, such as hydraulic fracturing or acid treatment (Dake L. P., 2009, [59]). In the case of a gas condensate field, periodic reverse injection of dry gas (cycling process) is sometimes used, as a result of which the reservoir pressure is maintained at a high level above the dew point pressure, which does not allow the gas to condense.

2.4.1. A dynamic model for the functioning of a gas deposit

Let's define a deposit as an underground accumulation of natural gas, all parts of which are hydrodynamically connected, i.e. development in one part of the deposit manifests itself in a change in reservoir pressure throughout the deposit. This hydrodynamic connection can extend for tens of kilometers or more, supported by groundwater (Zakirov S. N., Lapuk B. B., 1974, [60]).

When modelling the process of developing a gas deposit, the «average well» method is often used, in which it is assumed that all producing wells have the same characteristics, average for the deposit: average depth, average plume length, average allowable flow rate and depression, average coefficient of filtration resistance, etc. The reservoir pressure is also assumed to be equal to the average for the deposit.

When modelling a gas deposit, the following basic equations must be taken into account [60]:

1) material balance for a gas deposit;
2) the technological operation mode of the producing well;
3) gas inflow to the bottom of the well;
4) the relationship between the number of gas wells, production and flow rate of wells.

These equations underlie the Advanced Gas Production Planning System for a group of gas fields (AGPPS), which was actively used to carry out calculations for many groups of gas fields in the USSR (Margulov R. D., Khachaturov V. R., Fedoseev A. V., 1992, [61]). Each of these equations is described in detail below.

Equation of the material balance for a gas deposit

$$\frac{p(t)}{z(p(t))} = \frac{p_0 \cdot V(t)}{z(p_0) \cdot V_0} \tag{2.5}$$

describes the relationship between the average reservoir pressure $p(t)$ and the remaining gas reserve in the deposit $V(t)$. Here t is time, $z(p)$ is the coefficient of super-compressibility, taking into account the imperfection of natural gas, p_0 and V_0 are the initial values of reservoir pressure and gas reserve, respectively. It is worth noting that equation (2.5) does not take into account the displacement of gas by water, which occurs under the water pressure regime. We assume that the deposit is functioning in gas mode.

The remaining gas reserve $V(t)$ is associated with marginal production $Q(t)$ and total production $Q_{tot}(t)$:

$$Q_{tot}(t) = V_0 - V(t),$$

$$Q(t) = {Q_{tot}}'(t), \quad Q_{tot}(t) = \int\limits_0^t Q(t)dt. \tag{2.6}$$

The equations of the technological mode of operation of a producing well can be different, since the technological mode is set taking into account the action of limiting geological, technical, technological, environmental and other factors. In general, a technological mode is chosen in which the maximum possible flow rate is ensured while respecting the limiting factors. The following main technological modes can be noted: constant depression on the formation, constant flow rate, constant pressure gradient, constant downhole pressure, constant wellhead pressure, constant flow velocity. The constant debit mode is often used:

$$q(t) = q_0,$$

where $q(t)$ is the well flow rate, q_0 is the initial flow rate. It is assumed that the technological regime is the same for all producing wells, and the wells themselves are identical.

Equation of gas inflow to the bottom of the well

$$p^2(t) - p_{wf}^2(t) = A \cdot q(t) + B \cdot q^2(t) \tag{2.7}$$

connects the well flow rate $q(t)$, reservoir pressure $p(t)$ and bottom hole pressure $p_{wf}(t)$. Here A and B are the coefficients of filtration resistance, depending on the composition of the gas, the properties of the porous medium and other parameters.

The equation of the relationship between the number of gas wells, production and flow rate of wells. The last equation relates the number of producing wells $N(t)$, the marginal production $Q(t)$ and the flow rate $q(t)$:

$$N(t) = Q(t)/q(t). \tag{2.8}$$

Here it is assumed that $N(t)$ is a real number. This simplification is acceptable, since for large deposits, the number of wells can reach several tens or even hundreds by the period of constant production.

There are three stages in the development of a gas deposit: increasing, constant and falling production [57, 61]. At the stage of increasing production, production wells are being built, the field is being developed and production is reaching a constant level. At the end of this stage, the total gas production is 20-25% of the initial reserves. At the stage of continuous production, marginal production is maintained at a constant level due to the introduction of new wells and an increase in the capacity of booster compressor stations (BCS). The stage ends when further drilling of the deposit and increasing the capacity of the BCS becomes economically impractical. During the first two stages, the main reserves of the deposit are selected (about 60%). At the stage of declining production, the well stock usually remains unchanged, while the reservoir pressure drops significantly. The third stage ends when mining becomes unprofitable. Approximately 10-15% of the initial reserves remain untreated.

In the case of using the constant flow rate mode, the following scheme is usually used to calculate the main technological indicators of the first two stages [61]: first, the required number of wells $N(t)$ is calculated using (2.8) for a given marginal production $Q(t)$ and well flow rate q_0, then using (2.5-2.6), reservoir pressure $p(t)$ is found, after which the bottom-hole pressure $p_{wf}(t)$ is found according to (2.7), which is further used to determine the required capacities of the BCS.

At the stage of declining production, the number of producing wells is usually constant, so the marginal production decreases in proportion to the well flow rate. But since information about the coefficients of filtration resistances A and B and the operating mode at the last stage of development is very inaccurate or not known at all in advance, it is proposed to use the following empirical formula in the AGPPS to calculate the maximum production at this stage instead of (2.5-2.8):

$$Q(t) = \overline{Q} \cdot e^{-\frac{\overline{Q}}{\overline{V}}(t-\bar{t})}, \tag{2.9}$$

where $\overline{Q}$ is the maximum production at the second stage, $\bar{t}$ is the end time of the second stage, $\overline{V} = V(\bar{t})$ is the remaining gas reserve in the deposits at the end of the second stage. The formula (2.9) is used until the remaining gas reserve reaches the value $\overline{\overline{V}}$, after which production stops. It is assumed that $\overline{V} = \overline{\beta} \cdot V_0$, $\overline{\overline{V}} = \overline{\overline{\beta}} \cdot V_0$, where $\overline{\beta}$, $\overline{\overline{\beta}}$ are the set parameters. The $\overline{\beta}$ parameter actually depends on the marginal production at the second stage of $\overline{Q}$: the higher $\overline{Q}$, the lower $\overline{\beta}$. However, to simplify the model, this dependence is not taken into account.

In [61], an aggregated model of the deposit is proposed, for which the following additional assumptions are introduced:

- the wells are operated in a constant flow rate mode;
- the stage of falling production occurs at moment $\bar{t}$: $V(\bar{t}) = \overline{V}$;
- at the stage of falling production, the marginal production changes according to (2.9);
- the input of the BCS is not taken into account.

As a result, the behaviour of a gas deposit is described by the following system of differential equations:

$$\begin{cases} \dot{N} = n, \\ \dot{V} = -q \cdot N, \\ \dot{q} = \begin{cases} 0, & V > \overline{V}, \\ -\dfrac{q_0}{\overline{\overline{V}}} \cdot q \cdot N, & V \leq \overline{V}, \end{cases} \\ N(0) = 0,\ V(0) = V_0,\ q(0) = q_0. \end{cases}$$

Here, the rate of well entry $n(t) \geq 0$ is a control, and the following restriction must be met for it:

$$n(t) = 0, \text{ if } V(t) \leq \overline{V}.$$

The remaining values are calculated according to the following ratios:

$$\frac{p(t)}{z(p(t))} = \frac{p_0}{z(p_0)} \cdot \frac{V(t)}{V_0},$$

$$Q(t) = \begin{cases} q(t) \cdot N(t), & V > \overline{\overline{\overline{V}}}, \\ 0, & V \leq \overline{\overline{\overline{V}}}, \end{cases}$$

$$p_{wf}(t) = \sqrt{p^2(t) - A \cdot q(t) - B \cdot q^2(t)}.$$

Thus, marginal production $Q(t)$ depends on the rate of well entry $n(t)$:

$$Q(t) \equiv \widehat{Q}(n(\cdot), t). \tag{2.10}$$

In (2.10) $n(\cdot)$ is considered as a parameter function. By varying the $n(\cdot)$ function, different production dynamics can be obtained. Additional restrictions can be imposed on the management due to various technical reasons (for example, a limit on the rate of well entry: $n(t) \leq n_{\max}$).

The work [61] also provides for the construction of reserve wells, which, in addition to the reserve wells themselves, can include observation and other non-producing wells. It is assumed that the well reserve is created evenly during the period of continuous production. In this paper, we assume that the construction of reserve wells takes place simultaneously with producing wells, and the rate of input of reserve wells $n_p(t)$ is proportional to the rate of input of producing wells and is described by the reserve coefficient k_p:

$$n_p(t) = k_p \cdot n(t).$$

The reserve ratio can be assumed to be 10-20%. Let's denote via c_I capital investments in the construction of one well, and via c_e the cost of producing a unit of volume of gas. At the same time, we assume that construction costs are charged continuously, and marginal production costs do not depend on reservoir pressure, technological mode of well operation and other values (in particular, on the current capacity of BCS). As a result, the following function of marginal costs for the operation of a gas deposit is obtained:

$$c(n(\cdot), t) = c_I \cdot (1 + k_p) \cdot n(t) + c_e \cdot Q(n(\cdot), t).$$

Taking into account the coefficient δ - the difference between the interest rate on bank deposits and the inflation rate, the total costs given at the initial time are equal to the following value:

$$C(n(\cdot)) = \int_0^{+\infty} [c_I \cdot (1 + k_p) \cdot n(t) + c_e \cdot Q(n(\cdot), t)] e^{-\delta t} dt.$$

2.4.2. Simplified model for the functioning of a gas deposit

Now let's consider a simplified model of the functioning of a gas deposit. Let's introduce the following assumptions:

- the initial reserves of natural gas in the deposits are so large that at the considered planning interval, reducing gas reserves and reservoir pressure in the deposits are considered insignificant;
- all the wells being introduced have the same characteristics, the number of wells can be considered substantial, and the construction of wells takes place instantly.

Let's enter the following parameters:

- q_0 - initial well flow rate (t.c.f./year);
- c_e - the cost of producing a unit of gas volume (rub./t.c.f.);
- c_I - the specified capital investments in the construction of one well (rubles/year); reduction is carried out to an annual period of time, taking into account the expected service life of the well, the interest rate on bank deposits and the inflation rate (similar to paragraph 1.1.2. and (1.1));
- $k_p > 0$ is the reserve coefficient for well construction.

In this case, the flow rate of each well does not change over time, and the production cost function for a gas deposit is linear and has the following form:

$$\bar{c}(v) = \left((1 + k_p) \cdot \frac{c_I}{q_0} + c_e\right) \cdot v,$$

where $v \geq 0$ is the annual constant volume of natural gas production (t.c.f./year). The unit of measurement of the value of $\bar{c}(v)$ is rub./year.

2.4.3. Estimation of the production cost function

Let's build an estimate of the production cost function for a gas field according to a simplified model. Let's assume that the field consists of one large deposit. According to [81], as of June 1, 2016, the cost of producing 1 000 m^3 of natural gas from Gazprom was \$20 (1 320 rubles at the exchange rate on the same date). Thus, in prices at the beginning of 2022, the cost of production of 1 t.c.f. of natural gas is approximately 1 467 rubles.

To estimate the capital costs for the construction of wells, we use data on the construction of the Chayandinsky field (Yakutia). According to official sources [77, 82], the projected annual productivity of the field is 25 billion m^3 of natural gas, and as of April 27, 2018, the cost of developing the field was estimated at 450 billion rubles (543.3 billion rubles in prices at the beginning of 2022). Assuming a well life of 30 years and the equality of expected interest rates on bank deposits and inflation levels, we obtain the following estimate of production costs reduced to the annual period in prices at the beginning of 2022:

$$\bar{c}(v) = 2\ 095 \cdot v, \tag{2.11}$$

that is, the marginal production costs are 2 095 rubles./t.c.f.

2.5. Estimation of the demand function for natural gas in a non-gasified node

This paragraph proposes a method for estimating the gas demand function in a node that is an arbitrary non-gasified entity or municipality (MO) of the Russian Federation. Smaller units such as thermal power plants, boiler houses, boiler house groups, etc. can also act as nodes. The method under consideration is based on the division of the main gas consumers into several groups, for each of which an independent assessment of the demand function is carried out. At the same time, long-term changes in gas demand associated with the possible restructuring[24] of the electricity market in the node as a whole are not taken into account.

2.5.1. Overview of potential gas consumers on a non-gasified territories

During the gasification of the territory, the most economically significant transformations are the substitution of gas for other economically less profitable energy resources used in the territory and the development of the chemical industry using gas as a raw material. Let's highlight the following components among the main potential transformations:

- conversion of thermal power plants (TPP) and combined heat and power plants (CHP) to gas fuel;
- conversion of thermal boilers to gas fuel;
- conversion of enterprises to gas fuel;
- connection to the gas supply of the rural population;
- connection to the gas supply of the urban population (not included in the model);
- the development of the chemical industry (not taken into account in the model).

Transformations from this list imply the implementation of capital investments for the following purposes:

- laying of gas lines to the place of consumption;
- switching to new equipment (or upgrading an existing one) capable of working with a new type of fuel if gas is used as an energy source;
- construction and development of chemical enterprises, if gas is used as a raw material.

If gas is used as an energy source, then a full or partial transition to gas is carried out, and the fuel used before with the necessary equipment can remain as a backup. The expected economic benefit of such transformations is based on a reduction in fuel costs. When assessing the economic benefits, it is also possible to take into account the environmental component associated with a change in the amount of environmental pollution during the transition to a new type of fuel.

Components related to the connection to the gas supply of the urban population and the

[24] This restructuring may be a consequence of the gasification of the node.

development of the chemical industry, it was decided to exclude from consideration when assessing the demand function for natural gas for the following reasons. Connecting to the gas supply of the urban population is fraught with significant problems. The main problem is that, in fact, for multi-storey buildings located on a non-gasified territory, the possibility of connecting them to pipeline gas was not initially envisaged. Therefore, in order to prepare houses for connection, significant capital investments are needed to connect gas pipes to apartments, modernize the ventilation system and other necessary transformations, which are often economically impractical.

The second problem is related to security. As you know, gas is explosive and poses a serious threat to residents of apartment buildings. Everyone remembers the sad news from the media about explosions in residential buildings, as a result of which entire entrances collapsed, while human casualties were measured in dozens of people. Therefore, currently there is a tendency to move away from gas stoves in apartment buildings, new buildings are mostly built without connection to gas, even in gasified areas.

As for the chemical industry, there are currently about 15 large gas processing enterprises operating in Russia, providing over 90% of the total processing. Decisions on the construction of such enterprises and their location are usually made at the highest level, while the demand for the final products of such enterprises is often not determined by the needs of a particular region, therefore it is difficult to estimate this component of gas demand in advance.

2.5.2. Mathematical model for estimating the demand function for natural gas

In general, the demand for natural gas at node $i \in N$ is represented as a function $D_i(p_i, p_1^1, ..., p_i^K)$, where p_i is the price of natural gas at this node, and p_i^1, ..., p_i^K are prices for other energy sources. The following describes a somewhat simplified consumption model in which the demand for natural gas depends only on the price of it.

Let's build a mathematical model of consumption based on the transformations components introduced in paragraph 2.5.1., while not taking into account the components associated with connecting the urban population to gas and the development of the chemical industry. Let's split the total demand in node $i \in N$ in accordance with these components, in this case the demand function $D_i(p_i)$ is represented as the sum of demand functions for various components of the node:

- $D_i^{TPS}(p_i)$ - for thermal power plants and CHP;
- $D_i^{B}(p_i)$ - for thermal boilers;

- $D_i^{Ind}(p_i)$ - for enterprises;
- $D_i^{Vil}(p_i)$ - for the rural population.

Boiler units. Let's denote by R the set of all alternative fuels to natural gas, and by A the set of types of boiler units that are used by consumers and allow converting fuel into thermal energy. Each type $a \in A$ is defined by the following parameters:

- $r_a \in R$ is the type of fuel that the boiler uses;
- $V_a^{\max}$ is capacity (t.c.f./ year), i.e. the maximum amount of useful thermal energy generation in the case of maximum load; in the case of boilers for thermal power plants and CHP, capacity is understood as the maximum possible annual amount of energy transferred to the working steam;
- $\eta_a \in [0, 1]$ is the efficiency of the boiler unit, which is equal to the ratio of the generated useful thermal energy to the initial fuel energy.

Thermal power plants and CHP. In this paragraph, for brevity, we will call thermal power plants and CHP by plantsplant. Large plants should be separated into separate nodes or placed in the center of the node, and transportation costs for the gas pipeline sections suitable for them should be estimated using the function $\overline{E}^m(q, l)$ according to (2.1). Small plants can also be separated into separate nodes (in this case, the function $\overline{E}^{d,1}(q, l)$ is used to estimate the transport costs for the underwater section to this plant, if the plant is not an intermediate node between large nodes (see (2.3))), or taken into account as one of the components the node. In the latter case, the plant is connected to the «center» of the node by a section of the gas pipeline, and the function $\overline{E}^{d,2}(q, l)$ is used to estimate the corresponding transport costs (see (2.4)). The center of the node is understood as the geographical point at which the sections of gas pipelines connecting the node with adjacent nodes are connected, as well as from where the gas distribution network of the node comes out. When considering a plant as one of the components of a node, the accuracy of estimates decreases, but at the same time, the exhaustive complexity of the initial optimization problem also decreases. Let TPS_i be the set of active plants in node $i \in N$.

General model. Let the plant $s \in TPS_i$ be characterized by the following parameters:

- d_s is the current total heating capacity of the plant's boilers (t.c.f./year);
- $\zeta_s \geq 1$ is the reserve coefficient of the plant, equal to the ratio of the maximum possible daily heating capacity of the plant to the average daily one; this coefficient takes into account the unevenness in the output volumes of the plant associated with the influence of weather and time conditions on consumption volumes, as well as other factors;
- $n_{s,a}$ is the number of available boilers of type $a \in A$;
- $c_{s,r}$ is the cost of extraction and delivery of a unit of fuel $r \in R$ to the plant, (rub/t.c.f.); this

also includes all costs associated with fuel preparation: transportation, processing, storage;

- $c_s^{g,b}$ is the reduced capital costs for increasing the capacity of gas boilers per unit ((rub/year)/(t.c.f./year) = rub/t.c.f.);
- $\eta_s^g \in [0,1]$ is the efficiency of a gas boiler unit, which is equal to the ratio of the generated useful thermal energy to the initial fuel energy;
- l_s is the distance from the center of the node to the plant (km); if the plant is allocated to a separate node, then this value is zero.

Capital expenditures are adjusted to an annual period of time, taking into account the expected service life of the gas pipeline, the interest rate on bank deposits and the inflation rate (see paragraph 1.1.2. and (1.1)).

Let's denote by c_s^g the internal costs of using a unit of natural gas (rubles /t.c.f.), which include the following: the cost of transporting a unit of gas from the center of the node to the plant, the maximum reduced capital costs for laying the corresponding underwater section of the gas pipeline, the maximum reduced capital costs for preparing boilers, taking into account redundancy:

$$c_s^g = e^d \cdot l_s + \zeta_s \cdot c_s^{g,b}.$$

Denote by $a_1, ..., a_n$ the types of boilers available at the plant, sorted in order of increasing cost of producing a unit of heat $c_i = \frac{c_{s,r_{a_i}}}{\eta_{a_i}}$ (rubles/t.c.f.). In $V_i = \frac{n_{s,a_i} \cdot V_{a_i}^{\max}}{\zeta_s}$ (t.c.f./year), we denote the maximum amount of thermal energy generation using type a_i boilers available at the plant, taking into account redundancy. Let's consider the task of minimizing the cost of maintaining the total heating capacity d_s of the boilers of the plant. Let the demand function $D_s(p_i)$ determine the dependence of the optimal volume of natural gas consumed by the plant s (t.c.f./year) from the nodal price p_i for natural gas (rubles/t.c.f.) in terms of this problem. Let's denote

$$\widehat{D}_s(\widehat{p}) \equiv \max(0, \overline{D}_s(\widehat{p})), \tag{2.12}$$

where

$$\overline{D}_s(\widehat{p}) = \begin{cases} d_s, & \widehat{p} \in [0, c_1), \\ [d_s - V_1, d_s], & \widehat{p} = c_1, \\ d_s - V_1, & \widehat{p} \in (c_1, c_2), \\ [d_s - (V_1 + V_2), d_s - V_1], & \widehat{p} = c_2, \\ \dots & \\ [d_s - \sum\limits_{i=1}^{n} V_i, d_s - \sum\limits_{i=1}^{n-1} V_i], & \widehat{p} = c_n, \\ d_s - \sum\limits_{i=1}^{n} V_i, & \widehat{p} \in (c_n, +\infty). \end{cases} \tag{2.13}$$

Theorem 14. *The demand function $D_s(p_i)$ for natural gas from the plant s of node i is calculated as follows:*

$$D_s(p_i) = \frac{1}{\eta_s^g} \cdot \widehat{D}_s\left(\frac{p_i \ + \ c_s^g}{\eta_s^g}\right). \tag{2.14}$$

Proof. *According to the welfare theorem [7], the optimal output volumes of boilers are equal to the volumes corresponding to a competitive equilibrium in a model with a single price for the generated thermal energy, a consumer with a fixed consumption volume d_s and a producer who can dispose of the corresponding boilers. These boilers, together with the volume of heat consumed d_s form a piecewise constant function of the residual demand for thermal energy (2.12), where $\widehat{p}$ is the price of the generated thermal energy (rubles/t.c.f.).*

Without the use of natural gas, the equilibrium price $\widetilde{p}$ is equal to the minimum price $\widehat{p}$, satisfying the condition $\widehat{D}_s(\widehat{p}) = 0$. The benefit of using gas is equal to $\int\limits_{\frac{p_i \ + \ c_s^g}{\eta_s^g}}^{+\infty} \widehat{D}_s(\widehat{p})d\widehat{p}$, where p_i is the price of natural gas at node i (rub/t.c.f.). Here, the value $\frac{p_i \ + \ c_s^g}{\eta_s^g}$ is equal to the cost of producing thermal energy using natural gas (rub/t.c.f.). Thus, the optimal volume of thermal energy generated using natural gas is equal to $\widehat{D}_s\left(\frac{p_i \ + \ c_s^g}{\eta_s^g}\right)$, and the dependence $D_s(p_i)$ of the optimal volume of natural gas consumed (t.c.f./year) on the nodal price p_i is a piecewise constant function and is found using the ratio (2.14). ■

The graph of the function $\widehat{D}_s(\widehat{p})$ is shown in figure 30. The function $D_s(p_i)$ is obtained from the function $\widehat{D}_s(\widehat{p})$ by sequentially applying the following operations:

1) shifting the graph of the function to the left by c_s^g;
2) stretching the graph of the function along the abscissa axis by a factor of η_s^g;
3) narrowing the graph of the function along the ordinate axis by a factor of η_s^g.

The total demand in node i generated by all plants is equal to

$$D_i^{TPS}(p_i) = \sum_{s \in TPS_i} D_s(p_i).$$

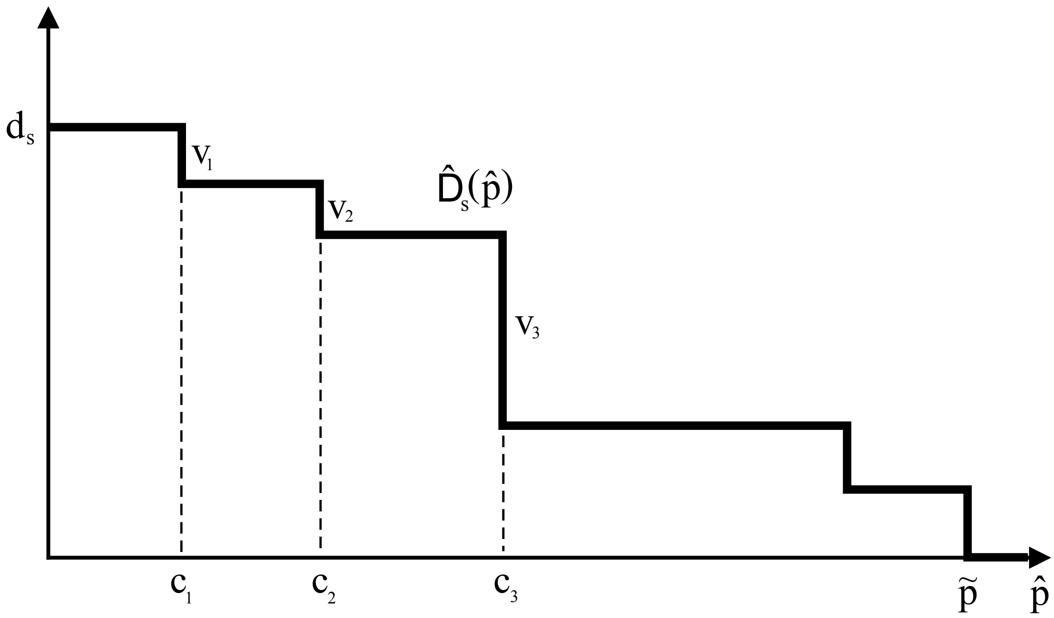

Fig. 30. Type of residual demand function $\widehat{D}_s(\widehat{p})$ for thermal power plants and CHP

Simplified model. Now let's consider a simplified model. Let's assume that the plant is allocated to a separate node, and it uses only one type of fuel. We introduce the following parameters characterizing the plant s:

- v_s - annual fuel consumption at the plant (t.c.f./year);
- c_s - the cost of extracting and delivering a unit of the fuel type used to the plant (similar to the general model, rub/t.c.f.).

Suppose that when switching to natural gas, the same boilers are used with a small upgrade, while maintaining the efficiency of the boilers, and the capital costs of modernization are insignificant and therefore not taken into account. The process of converting solid fuel boilers to natural gas is described, for example, in [62].

For such a simplified model, switching to natural gas becomes beneficial if the cost of buying a unit of gas is less than the cost of producing and delivering a unit of currently used fuel. Thus, the natural gas demand function for plant s has the following form:

$$D_s(p_i) = \begin{cases} v_s, & p_i \in [0, c_s), \\ [0, v_s], & p_i = c_s, \\ 0, & p_i \in (c_s, +\infty), \end{cases} \tag{2.15}$$

where p_i is the price of natural gas at node i (rub/t.c.f.). Since the plant is allocated to a separate node, the total demand $D_i^{TPS}(p_i)$ generated by all plants coincides with $D_s(p_i)$.

Thermal boilers. The process of converting boiler houses to natural gas is generally similar to the process of converting thermal power plants and CHP. The only significant difference is that the capacity of boiler units in boiler houses, as a rule, is much less than similar capacities for thermal power plants and CHP. Denote by B_i the set of operating boiler houses in node $i \in N$.

General model. In general, the parameters characterizing the boiler house are similar to those characterizing a thermal power plant or CHP, therefore, the demand for gas from the thermal boiler house is determined similarly to the previous paragraph. Let's denote the demand function of boiler house $b \in B_i$ by $D_b(p_i)$. The total demand in node i generated by all plants is equal to

$$D_i^B(p_i) = \sum_{b \in B_i} D_b(p_i).$$

Simplified model. Now let's consider a simplified model. Let node i under consideration be a territory or a municipality characterized by the following parameters:

- v_i^b - total fuel consumption by boiler houses (t.c.f./year);
- c_i^b - the average cost of extraction and delivery to the boiler house of a unit of the type of fuel used (rub/t.c.f.);
- S_i^b - the area of the node's territory (km^2).

Let's make the following assumptions:

- boiler houses are characterized by approximately the same fuel consumption, boiler efficiency and the cost of extraction, delivery and preparation of a unit of fuel;
- when converting boilers to natural gas, the same boilers are used with a small upgrade, while maintaining the efficiency of boilers, and the capital costs of modernization are insignificant and therefore not taken into account.

To estimate the demand function $D_i^B(p_i)$, it is necessary to take into account the transport costs associated with the construction and operation of the gas distribution network connecting the center of the node with the boiler houses. To estimate transport costs, we use the formula

(2.4).

A node in the shape of a circle. Let's consider an idealized case in which the territory of the node has the shape of a circle, the center of the node coincides with the center of the circle, and an infinite number of boiler houses are distributed evenly over the territory of the node[25].

Theorem 15. *The demand function $D_i^B(p_i)$ for natural gas from the rural population of node i for this idealized case is calculated as follows. If $c_i^b < e^d\sqrt{\frac{S_i^b}{\pi}}$, then*

$$D_i^B(p_i) = \begin{cases} \frac{\pi \cdot v_i^b}{S_i^b \cdot {e^d}^2}\left(c_i^b - p_i\right)^2, & p_i \in [0, c_i^b], \\ 0, & p_i \in (c_i^b, +\infty). \end{cases} \tag{2.16}$$

If $c_i^b \geq e^d\sqrt{\frac{S_i^b}{\pi}}$, then

$$D_i^B(p_i) = \begin{cases} v_i^b, & p_i \in \left[0, c_i^b - e^d\sqrt{\frac{S_i^b}{\pi}}\right], \\ \frac{\pi \cdot v_i^b}{S_i^b \cdot {e^d}^2}\left(c_i^b - p_i\right)^2, & p_i \in \left(c_i^b - e^d\sqrt{\frac{S_i^b}{\pi}}, c_i^b\right], \\ 0, & p_i \in (c_i^b, +\infty). \end{cases} \tag{2.17}$$

Proof. *The radius of the circle representing the territory of the node is equal to $\sqrt{\frac{S_i^b}{\pi}}$. For boilers located at a distance $r \in [0, \sqrt{\frac{S_i^b}{\pi}}]$ from the center of the node, the marginal cost of heat generation for the available fuel is equal to c_i^b, and in the case of using natural gas - $(p_i + e^d \cdot r)$, where p_i is the price of natural gas in node i (rub/t.c.f.). It follows that switching to gas is beneficial if $r \leq \frac{c_i^b - p_i}{e^d}$. Consider a strip of boiler houses of infinitely small width Δr, located at a distance r from the center of the node (figure 31). Its area is equal to $2\pi \cdot r \cdot \Delta r$, which means that the fuel consumption of the boilers of this strip is $v_i^b \cdot \frac{2\pi \cdot r \cdot \Delta r}{S_i^b}$ (t.c.f./year).Thus, the demand for natural gas from the boiler houses of the node is calculated as follows:*

$$D_i^B(p_i) = \int\limits_0^{\min\left(\max\left(\frac{c_i^b - p_i}{e^d}, 0\right), \sqrt{\frac{S_i^b}{\pi}}\right)} v_i^b \cdot \frac{2\pi \cdot r}{S_i^b} dr = \frac{\pi \cdot v_i^b}{S_i^b}\left(\min\left(\max\left(\frac{c_i^b - p_i}{e^d}, 0\right), \sqrt{\frac{S_i^b}{\pi}}\right)\right)^2.$$

[25] This assumption is highly simplifying, but it is acceptable, taking into account the absence of a fixed component in the used estimate of the transport cost function for gas pipelines (2.4) (the estimate of the gas demand function for evenly dispersed boiler houses in a small area with total fuel consumption v is approximated by the estimate of the gas demand function for one boiler house with fuel consumption v located in the center in this area).

The latter ratio implies fairness (2.16, 2.17). ■

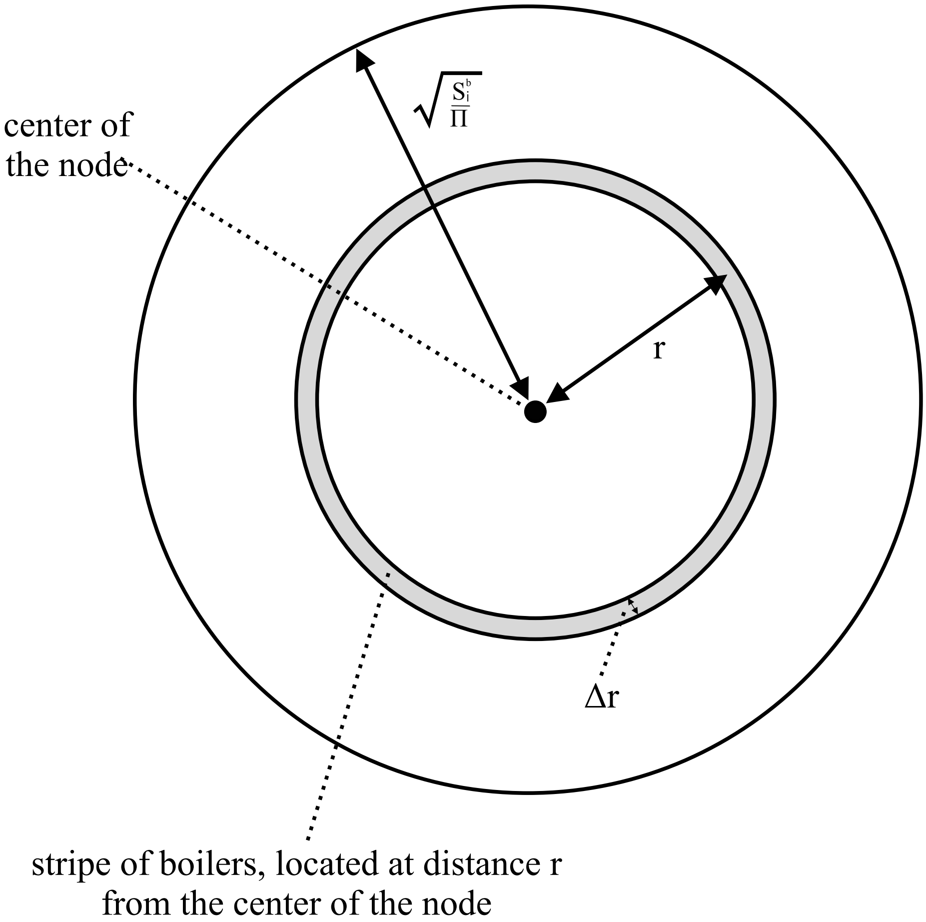

Fig. 31. The territory of the node is in the form of a circle with an infinite number of boiler houses distributed evenly over the territory of the node

The estimates obtained (2.16, 2.17) can also be applied to an arbitrary shape of the site and any number of boiler houses as a simplification.

A node in the shape of a circle. Now let's consider the case in which the territory of the node has the shape of a rectangle of length w and width S_i^b/w. Suppose that the center of the node coincides with the center of the rectangle, and an infinite number of boiler houses are distributed evenly over the territory of the node, similar to the case of a node in the shape of a circle. Let's denote

$$R(p_i) \equiv \frac{c_i^b - p_i}{e^d}.$$

Theorem 16. *The demand function $D_i^B(p_i)$ for natural gas from the boiler houses of node i or this case is calculated as follows. If $R(p_i) \leq 0$, then $D_i^B(p_i) = 0$. Otherwise, the following cases are possible.*

1. If $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \leq R(p_i)^2$, then

$$D_i^B(p_i) = v_i^b. \tag{2.18}$$

2. If $w/2 \geq R(p_i)$, $\dfrac{S_i^b}{2w} \geq R(p_i)$, *then*

$$D_i^B(p_i) = \frac{\pi \cdot v_i^b \cdot R(p_i)^2}{S_i^b}. \tag{2.19}$$

3. If $w/2 \leq R(p_i)$, $\dfrac{S_i^b}{2w} \geq R(p_i)$, *then*

$$D_i^B(p_i) = \frac{v_i^b}{S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right]. \tag{2.20}$$

4. If $(w/2)^2 + \left(\dfrac{S_i^b}{2w}\right)^2 \geq R(p_i)^2$, $w/2 \leq R(p_i)$, $\dfrac{S_i^b}{2w} \leq R(p_i)$, *then*

$$D_i^B(p_i) = \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{v_i^b}{S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right] - \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arccos\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right). \tag{2.21}$$

5. If $w/2 \geq R(p_i)$, $\dfrac{S_i^b}{2w} \leq R(p_i)$, *then*

$$D_i^B(p_i) = \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arcsin\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right). \tag{2.22}$$

Proof. *Similarly to the case of a node in the form of a circle, it is shown that the conversion of an arbitrary boiler unit to gas is advantageous if the distance from it to the center of the node does not exceed* $\dfrac{c_i^b - p_i}{e^d}$, *where* p_i *is the price of natural gas in node* i *(rub/t.c.f.). Thus, the intersection of a circle of a radius* $R(p_i) \equiv \dfrac{c_i^b - p_i}{e^d}$ *centered in the center of the node and a rectangle of length* w *and width* S_i^b/w *centered in the center of the node defines the area for those boilers that benefit from switching to gas. Figure 32 shows this area on a coordinate system with the origin in the center of the node. The part of the boundary of this circle belonging to the first quarter of the coordinate plane is described by the ratio* $y = \sqrt{R(p_i)^2 - x^2}$.

Consider a rectangular boiler house area of infinitesimal width Δx *and length* Δy *belonging to a node. Its area is equal to* $\Delta x \cdot \Delta y$, *which means that the fuel consumption of boilers in this area is* $v_i^b \cdot \dfrac{\Delta x \cdot \Delta y}{S_i^b}$ *(t.c.f./year). Thus, the demand for natural gas from the boiler houses of the*

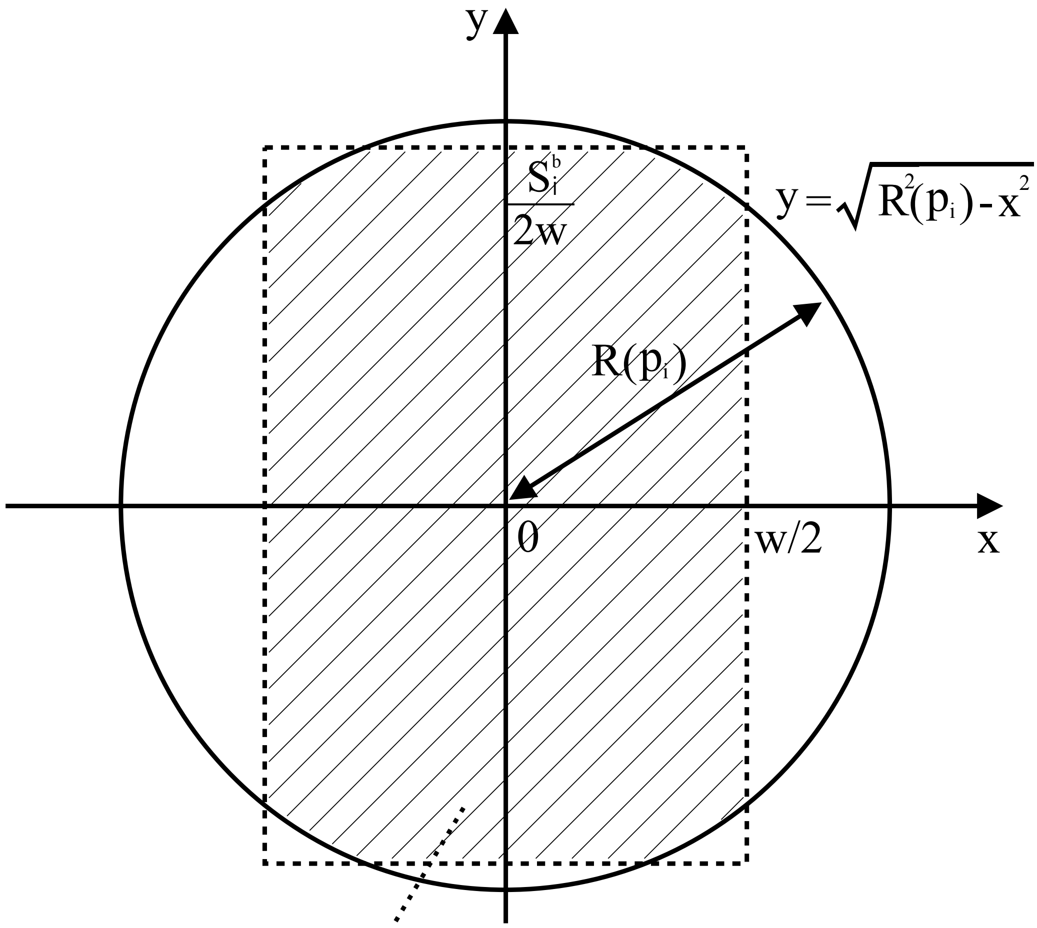

Fig. 32. The territory of the node is in the form of a circle with an infinite number of boiler houses distributed evenly over the territory of the node

node is calculated as follows:

$$D_i^B(p_i) = 4 \cdot \int\limits_0^{\min(w/2,\, R(p_i))} \left[\int\limits_0^{\min\left(\frac{S_i^b}{2w},\, \sqrt{R(p_i)^2 - x^2}\right)} \frac{v_i^b \cdot dy}{S_i^b} \right] dx$$

$$= \frac{4 \cdot v_i^b}{S_i^b} \cdot \int\limits_0^{\min(w/2,\, R(p_i))} \min\left(\frac{S_i^b}{2w},\, \sqrt{R(p_i)^2 - x^2}\right) dx.$$

Let's consider five cases of the relative position of the rectangle characterizing the territory of the node and the circle defining the area for those boiler houses that benefit from switching to gas (figure 33).

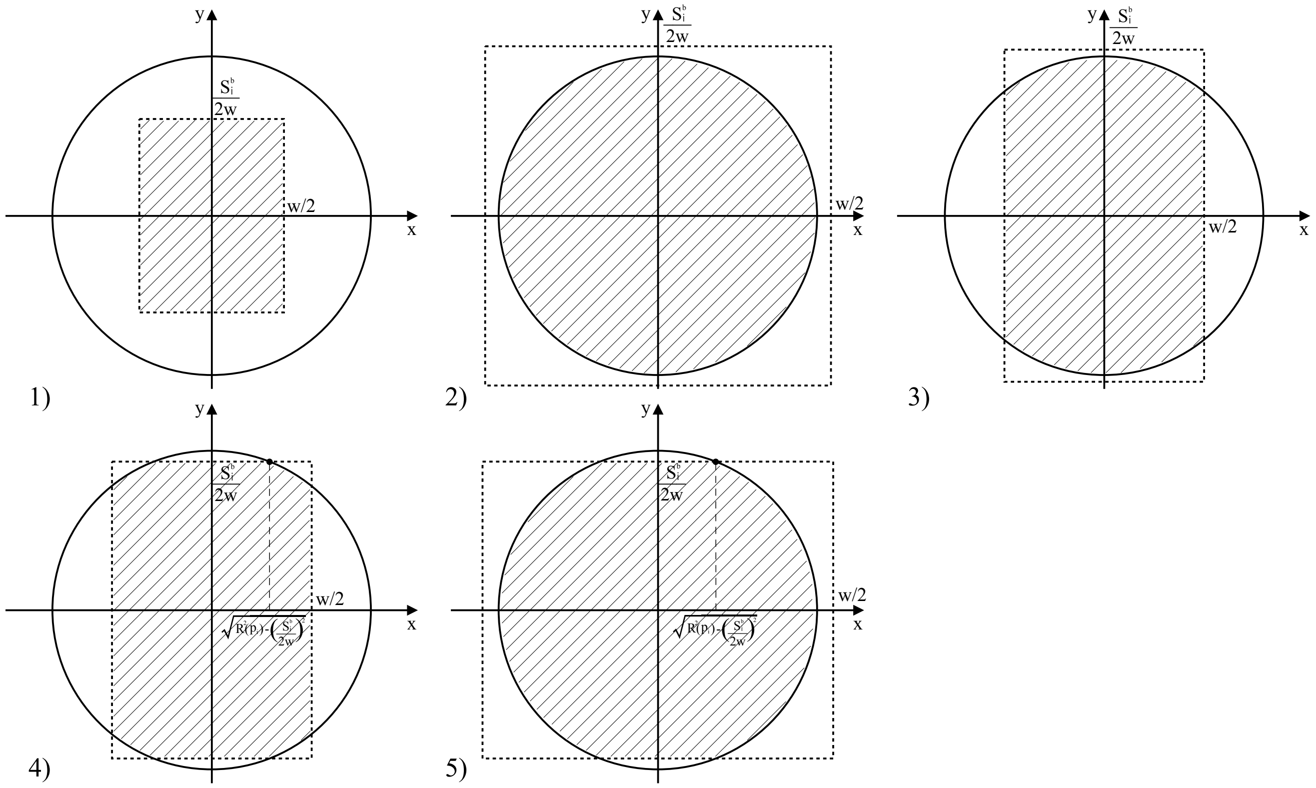

Fig. 33. Various cases of the relative position of the rectangle characterizing the territory of the node and the circle defining the area for those boilers that benefit switching to gas: 1) $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \leq R(p_i)^2$; 2) $w/2 \geq R(p_i)$, $\frac{S_i^b}{2w} \geq R(p_i)$; 3) $w/2 \leq R(p_i)$, $\frac{S_i^b}{2w} \geq R(p_i)$; 4) $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \geq R(p_i)^2$, $w/2 \leq R(p_i)$, $\frac{S_i^b}{2w} \leq R(p_i)$; 5) $w/2 \geq R(p_i)$, $\frac{S_i^b}{2w} \leq R(p_i)$

Case 1: $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \leq R(p_i)^2$. *In this case, the rectangle is entirely within the circle and* $D_i^B(p_i) = \frac{4 \cdot v_i^b}{S_i^b} \cdot \int\limits_0^{w/2} \frac{S_i^b \cdot dx}{2w} = \frac{4 \cdot v_i^b}{S_i^b} \cdot \frac{S_i^b}{2w} \cdot w/2 = v_i^b$, *that is, all boiler houses benefit from switching to gas.*

Case 2: $w/2 \geq R(p_i)$, $\frac{S_i^b}{2w} \geq R(p_i)$. *In this case*

$$D_i^B(p_i) = \frac{4 \cdot v_i^b}{S_i^b} \cdot \int\limits_0^{R(p_i)} \sqrt{R(p_i)^2 - x^2} \cdot dx$$

$$= \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{x}{2} \cdot \sqrt{R(p_i)^2 - x^2} + \frac{R(p_i)^2}{2} \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=0}^{R(p_i)}$$

$$= \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{R(p_i)^2}{2} \cdot \frac{\pi}{2}\right] = \frac{\pi \cdot v_i^b \cdot R(p_i)^2}{S_i^b}.$$

Case 3: *$w/2 \le R(p_i)$, $\dfrac{S_i^b}{2w} \ge R(p_i)$. In this case*

$$D_i^B(p_i) = \frac{4 \cdot v_i^b}{S_i^b} \cdot \int_0^{w/2} \sqrt{R(p_i)^2 - x^2} \cdot dx$$

$$= \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{x}{2} \cdot \sqrt{R(p_i)^2 - x^2} + \frac{R(p_i)^2}{2} \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=0}^{w/2}$$

$$= \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{w}{4} \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + \frac{R(p_i)^2}{2} \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right]$$

$$= \frac{v_i^b}{S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right].$$

Case 4: *$(w/2)^2 + \left(\dfrac{S_i^b}{2w}\right)^2 \ge R(p_i)^2$, $w/2 \le R(p_i)$, $\dfrac{S_i^b}{2w} \le R(p_i)$. In this case*

$$D_i^B(p_i) = \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\int_0^{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}} \frac{S_i^b}{2w} \cdot dx + \int_{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{w/2} \sqrt{R(p_i)^2 - x^2} \cdot dx \right]$$

$$= \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}$$

$$+ \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{x}{2} \cdot \sqrt{R(p_i)^2 - x^2} + \frac{R(p_i)^2}{2} \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{w/2}$$

$$= \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}$$

$$+ \frac{2 \cdot v_i^b}{S_i^b} \cdot \left[x \cdot \sqrt{R(p_i)^2 - x^2} + R(p_i)^2 \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{w/2} = \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}$$

$$+ \frac{v_i^b}{S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right]$$

$$- \frac{2 \cdot v_i^b}{S_i^b} \cdot \left[\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} \cdot \frac{S_i^b}{2w} + R(p_i)^2 \cdot \arcsin\left(\frac{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}{R(p_i)}\right)\right]$$

$$= \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{v_i^b}{S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right]$$

$$- \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arccos\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right).$$

Case 5: *$w/2 \geq R(p_i)$, $\frac{S_i^b}{2w} \leq R(p_i)$. In this case*

$$D_i^B(p_i) = \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\int\limits_0^{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}} \frac{S_i^b}{2w} \cdot dx + \int\limits_{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{R(p_i)} \sqrt{R(p_i)^2 - x^2} \cdot dx \right]$$

$$= \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}$$

$$+ \frac{4 \cdot v_i^b}{S_i^b} \cdot \left[\frac{x}{2} \cdot \sqrt{R(p_i)^2 - x^2} + \frac{R(p_i)^2}{2} \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{R(p_i)}$$

$$= \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}$$

$$+ \frac{2 \cdot v_i^b}{S_i^b} \cdot \left[x \cdot \sqrt{R(p_i)^2 - x^2} + R(p_i)^2 \cdot \arcsin\left(\frac{x}{R(p_i)}\right)\right]\Bigg|_{x=\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}^{R(p_i)}$$

$$= \frac{2 \cdot v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{\pi \cdot v_i^b \cdot R(p_i)^2}{S_i^b}$$

$$- \frac{2 \cdot v_i^b}{S_i^b} \cdot \left[\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} \cdot \frac{S_i^b}{2w} + R(p_i)^2 \cdot \arcsin\left(\frac{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}{R(p_i)}\right)\right]$$

$$= \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{\pi \cdot v_i^b \cdot R(p_i)^2}{S_i^b} - \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arcsin\left(\frac{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}{R(p_i)}\right)$$

$$= \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arccos\left(\frac{\sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2}}{R(p_i)}\right)$$

$$= \frac{v_i^b}{w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{2 \cdot v_i^b \cdot R(p_i)^2}{S_i^b} \cdot \arcsin\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right).$$

Validity of (2.18-2.22) follows from the cases considered. ■

Enterprises. Large enterprises with high fuel consumption should be separated into separate nodes or placed in the center of the node. Medium and small enterprises can be accounted for either as a separate node, or as one of the components of the node (similar to the case of thermal power plants and CHP). Let Ind_i be the set of enterprises in node $i \in N$ that require thermal energy to produce their own goods.

General model. Let enterprise $m \in Ind_i$ be characterized by the following parameters:

- p_m - the external price of the goods produced by the enterprise (rubles/unit);
- v_m - the amount of thermal energy consumed, which ensures the output of a unit of goods (t.c.f./unit);
- c_m - non-thermal costs in the production of a unit of goods (rub./unit);
- $V_m^{\max}$ - capacity, i.e. the maximum output volume (units/year);
- $n_{m,a}$, $c_{m,r}$, $c_m^{g,b}$, η_m^g, l_m - similar to thermal power plants and CHP.

Let's denote by c_m^g the internal costs of using a unit of natural gas (rub. / t.c.f.), which include the following: the cost of transporting a unit of gas from the center of the node to the plant, the maximum reduced capital costs for laying the corresponding underwater section of the gas pipeline, the maximum reduced capital costs for preparing boilers:

$$c_m^g = e^d \cdot l_m + c_m^{g,b}.$$

Let's consider the task of maximizing the profit of enterprise m. Suppose that $c_m < p_m$ (otherwise, it is not profitable to produce goods at any nodal gas price, i.e. there is no demand for gas). As for the case of thermal power plants and CHP, we will arrange the boilers $a_1, ..., a_n$ at the disposal of the enterprise by non-decreasing costs of producing a unit of heat $c_i \equiv \frac{c_{m,r_{a_i}}}{\eta_{a_i}}$ (rub./t.c.f.). Let's denote by $V_i \equiv n_{m,a_i} \cdot V_{a_i}^{\max}$ (t.c.f./year) the maximum volume of thermal energy generation using the type a_i boilers available at the enterprise. Let the demand function $D_m(p_i)$ determine the dependence of the optimal volume of natural gas consumed by enterprise m (t.c.f./year) from the nodal price p_i for natural gas (rub./t.c.f.) in terms of this task.

Let's denote

$$\widehat{D}_m(\widehat{p}) \equiv \begin{cases} \overline{D}_m(\widehat{p}), & \text{if } \overline{D}_m(\widehat{p}) \geq 0 \text{ and } \widehat{p} \leq \frac{p_m - c_m}{v_m}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.23}$$

where $\overline{D}_m(\widehat{p})$ is defined similarly to (2.13) with substitution d_s to $v_m \cdot V_m^{\max}$ (figure 34).

Theorem 17. *The demand function $D_m(p_i)$ for natural gas from enterprise m of node i is calculated as follows:*

$$D_m(p_i) = \frac{1}{\eta_m^g} \cdot \widehat{D}_m\left(\frac{p_i + c_m^g}{\eta_m^g}\right).$$

Proof. *The proof of this theorem is generally similar to the proof of theorem 14. It is only necessary to take into account that*

1) the piecewise constant function of the residual demand for thermal energy for this case has the form (2.23);

2) the value $\frac{p_m - c_m}{v_m}$ (rub./t.c.f.) is equal to the price of thermal energy, at which it becomes unprofitable for the enterprise to produce goods;

3) the value $v_m \cdot V_m^{\max}$ (t.c.f./year) is equal to the amount of thermal energy required to fully load production.

■

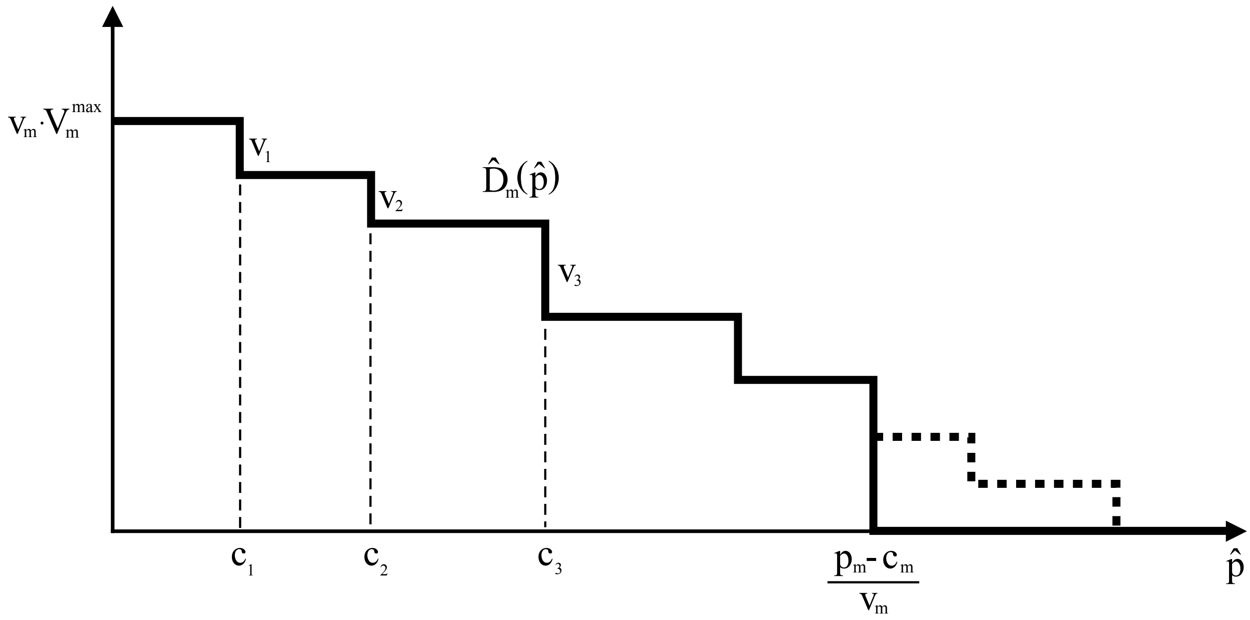

Fig. 34. Type of residual demand function $\widehat{D}_m(\widehat{p})$ for enterprises

Total demand in node i generated by all plants is equal to

$$D_i^{Ind}(p_i) = \sum_{m \in Ind_i} D_m(p_i).$$

Simplified model. The simplified model for estimating demand for an enterprise is completely identical to the similar model for thermal power plants and CHP. We assume that the enterprise $m \in Ind_i$ is allocated to a separate node and is set by the parameters v_m and c_m - annual fuel consumption (t.c.f./year) and the cost of extraction and delivery of a fuel unit used (rub./t.c.f.), respectively. The natural gas demand function for enterprise m has the following form:

$$D_m(p_i) = \begin{cases} v_m, & p_i \in [0, c_m), \\ [0, v_m], & p_i = c_m, \\ 0, & p_i \in (c_m, +\infty), \end{cases}$$

where p_i is the price of natural gas at node i (rub./t.c.f.). The total demand $D_i^{Ind}(p_i)$ generated by all enterprises of the node coincides with $D_m(p_i)$.

Rural population. Rural settlements should be separated into separate nodes. We assume that for each locality it is possible to connect residential buildings to gas for heating and cooking purposes, and all residential buildings have approximately the same characteristics and are evenly distributed throughout the node. In the case of gasification, an apartment building (in which, presumably, one family of h^{Vil} people live) is supplied with a gas boiler, the characteristics of which are common to all nodes: $C^{Vil,br}$ (rub./year) - the cost reduced to the annual period (reduction is carried out similarly to paragraph 1.1.2.), $\eta^{Vil,br}$ - efficiency. During the gasification of a settlement, we assume that a distribution network is being built inside the settlement, while each residential building is connected by a gas pipeline to the center of the node[26].

We assume that the rural population of node $i \in N$ is characterized by the following parameters:

- v_i^{Vil} - the current total final volume of thermal energy consumed by the settlement (t.c.f./year);
- η_i^{Vil} - average efficiency of available fuel combustion equipment;
- c_i^{Vil} - average current marginal costs for the extraction and delivery of fuel necessary for the generation of thermal energy (rub/t.c.f.);
- S_i^{Vil} - the area of the node's territory (km^2).
- G_i^{Vil} - population size.

To estimate the transport costs during the construction and operation of the distribution network, we use the formula (2.4).

A node in the shape of a circle. Let's consider an idealized case, as in the simplified model for estimating demand from boiler houses, in which the territory of the node has the shape of a circle, the center of the node coincides with the geographical center of the circle, and an infinite number of residential buildings are distributed evenly throughout the territory of the node.

Let's denote

$$\bar{c}_i^{Vil} = \eta^{Vil,br}\left(\frac{c_i^{Vil}}{\eta_i^{Vil}} - c_i^{Vil,br}\right), \tag{2.24}$$

where

$$c_i^{Vil,br} = \frac{C^{Vil,br} \cdot G_i^{Vil}}{v_i^{Vil} \cdot h^{Vil}}. \tag{2.25}$$

[26] The assumption that each residential building is connected to the center of the node by a gas pipeline is highly simplifying, but it is acceptable, taking into account the absence of a fixed component in the used estimate of the transport cost function for gas pipelines (2.4) (the estimate of transport costs for laying K gas pipelines for K residential houses located close to each other is approximated by the estimate of transport costs for laying one common gas pipeline used by all these houses).

Theorem 18. *The demand function $D_i^{Vil}(p_i)$ for natural gas from the rural population of node i for this idealized case is calculated as follows.*
If $\dfrac{c_i^{Vil}}{\eta_i^{Vil}} < c_i^{Vil,br}$, *then*

$$D_i^{Vil}(p_i) = 0. \tag{2.26}$$

If $0 \le \overline{c}_i^{Vil} < e^d\sqrt{\dfrac{S_i^{Vil}}{\pi}}$, *then*

$$D_i^{Vil}(p_i) = \frac{1}{\eta^{Vil,br}} \cdot \begin{cases} \dfrac{\pi \cdot v_i^{Vil}}{S_i^{Vil} \cdot e^{d2}} \left(\overline{c}_i^{Vil} - p_i\right)^2, & p_i \in \left[0, \overline{c}_i^{Vil}\right], \\ 0, & p_i \in \left(\overline{c}_i^{Vil}, +\infty\right). \end{cases} \tag{2.27}$$

If $\overline{c}_i^{Vil} \ge e^d\sqrt{\dfrac{S_i^{Vil}}{\pi}}$, *then*

$$D_i^{Vil}(p_i) = \frac{1}{\eta^{Vil,br}} \cdot \begin{cases} v_i^{Vil}, & p_i \in \left[0, \overline{c}_i^{Vil} - e^d\sqrt{\frac{S_i^{Vil}}{\pi}}\right], \\ \dfrac{\pi \cdot v_i^{Vil}}{S_i^{Vil} \cdot e^{d2}} \left(\overline{c}_i^{Vil} - p_i\right)^2, & p_i \in \left(\overline{c}_i^{Vil} - e^d\sqrt{\frac{S_i^{Vil}}{\pi}}, \overline{c}_i^{Vil}\right], \\ 0, & p_i \in \left(\overline{c}_i^{Vil}, +\infty\right). \end{cases} \tag{2.28}$$

Proof. *The number of residential buildings in the node is equal to* $\dfrac{G_i^{Vil}}{h^{Vil}}$, *therefore, with full gasification of the settlement, the total costs for the purchase of gas boilers are equal to* $\dfrac{C^{Vil,br} \cdot G_i^{Vil}}{h^{Vil}}$ *(rub./year), and the cost of a gas boiler in recalculation per unit of generated heat is expressed by the value (2.25).*

When applying reasoning similar to that used in estimating demand for thermal boilers, taking into account the cost of gas boilers and the population, the following formula is derived to estimate the demand function for natural gas from the rural population:

$$D_i^{Vil}(p_i) = \frac{1}{\eta^{Vil,br}} \cdot \frac{\pi \cdot v_i^{Vil}}{S_i^{Vil}} \left(\min\left(\max\left(\frac{\eta^{Vil,br}\left(\frac{c_i^{Vil}}{\eta_i^{Vil}} - c_i^{Vil,br}\right) - p_i}{e^d}, 0 \right), \sqrt{\frac{S_i^{Vil}}{\pi}} \right)\right)^2 .$$

The latter ratio implies that (2.26-2.28) hold. ■

The obtained estimates (2.26-2.28) can also be applied to an arbitrary shape of the site and any number of boiler houses as a simplification.

A node in the shape of a circle. Now let's consider the case in which the territory of the node has the shape of a rectangle of length w and width S_i^b/w. Suppose that the center of the node coincides with the center of the rectangle, and an infinite number of boiler houses are distributed evenly over the territory of the node, similar to the case of a node in the shape of a circle. Let's denote

$$R(p_i) \equiv \frac{\overline{c}_i^{Vil} - p_i}{e^d},$$

where $\overline{c}_i^{Vil}$ is defined similarly (2.24).

Theorem 19. *The demand function $D_i^{Vil}(p_i)$ for natural gas from the rural population of node i for this case is calculated as follows. If $R(p_i) \le 0$, then $D_i^{Vil}(p_i) = 0$. Otherwise, the following cases are possible.*

1. If $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \le R(p_i)^2$, then

$$D_i^{Vil}(p_i) = \frac{v_i^b}{\eta^{Vil,br}}. \tag{2.29}$$

2. If $w/2 \ge R(p_i)$, $\frac{S_i^b}{2w} \ge R(p_i)$, then

$$D_i^{Vil}(p_i) = \frac{\pi \cdot v_i^b \cdot R(p_i)^2}{\eta^{Vil,br} \cdot S_i^b}. \tag{2.30}$$

3. If $w/2 \le R(p_i)$, $\frac{S_i^b}{2w} \ge R(p_i)$, then

$$D_i^{Vil}(p_i) = \frac{v_i^b}{\eta^{Vil,br} \cdot S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right]. \tag{2.31}$$

4. If $(w/2)^2 + \left(\frac{S_i^b}{2w}\right)^2 \ge R(p_i)^2$, $w/2 \le R(p_i)$, $\frac{S_i^b}{2w} \le R(p_i)$, then

$$\begin{aligned} D_i^{Vil}(p_i) = {} & \frac{v_i^b}{\eta^{Vil,br} \cdot w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} \\ & + \frac{v_i^b}{\eta^{Vil,br} \cdot S_i^b} \cdot \left[w \cdot \sqrt{R(p_i)^2 - \frac{w^2}{4}} + 2 \cdot R(p_i)^2 \cdot \arcsin\left(\frac{w}{2 \cdot R(p_i)}\right)\right] \\ & - \frac{2 \cdot v_i^b \cdot R(p_i)^2}{\eta^{Vil,br} \cdot S_i^b} \cdot \arccos\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right). \end{aligned} \tag{2.32}$$

5. If $w/2 \geq R(p_i)$, $\frac{S_i^b}{2w} \leq R(p_i)$, *then*

$$D_i^{Vil}(p_i) = \frac{v_i^b}{\eta^{Vil,br} \cdot w} \cdot \sqrt{R(p_i)^2 - \left(\frac{S_i^b}{2w}\right)^2} + \frac{2 \cdot v_i^b \cdot R(p_i)^2}{\eta^{Vil,br} \cdot S_i^b} \cdot \arcsin\left(\frac{S_i^b}{2 \cdot w \cdot R(p_i)}\right). \quad (2.33)$$

The proof of the theorem is based on the application of reasoning similar to those used in the proofs of theorems 16 and 18.

It is worth noting that this model does not assume the presence of centralized heating in rural settlements. If there is one, then the corresponding demand for centralized boilers should be considered according to the model described in the subparagraph «thermal boilers». At the same time, when calculating demand related to individually heated houses, the parameters v_i^{Vil} and G_i^{Vil} described above should be adjusted, removing from consideration the part of the population using centralized heating.

Total demand and utility of consumption. The gas demand function in the node is equal to the sum of the components obtained:

$$D_i(p_i) = D_i^{TPS}(p_i) + D_i^B(p_i) + D_i^{Ind}(p_i) + D_i^{Vil}(p_i).$$

The consumption utility function $U_i(v_i^d)$ is related to the total demand function $D_i(p_i)$ by the ratios (1.11).

2.6. Analysis of gasification prospects in Irkutsk Oblast

Despite the obvious advantages of natural gas over other hydrocarbons, many regions of Russia are still not gasified. The rapid development of the domestic gas sector in the last decade, the experience gained in the implementation of such large-scale gas transportation projects as the «Power of Siberia» and «Nord Stream 2», as well as the refusal of some Western countries from Russian gas, give reason to consider the possibility of connecting new regions and cities of our country to pipeline gas. In this regard, the task of identifying the most promising schemes for the development of the Russian gas transmission network becomes urgent. To solve this problem, it is necessary to be able to assess the economic potential of gasification in specific areas of the country. This paragraph evaluates the prospects for gasification of Irkutsk Oblast. Most of the data used in it is taken from the decrees of the governor of Irkutsk Oblast on the approval of schemes and programmes for the development of the electric power industry [2–4].

Irkutsk Oblast is located on the territory of 767.9 thousand km^2 with a population of 2 357[27] thousand people. The population is characterized by a low density (3.07 people/km^2 with a national average of 8.58 people./km^2) and extremely uneven distribution. The southern and southwestern parts of the region are most densely populated along the banks of the Angara River and along the Trans-Siberian Railway. The urban population is 73%.

Water resources, among which it is worth highlighting the large rivers Lena, Angara, Nizhny Tunguska, as well as Lake Baikal, provide the region with fresh water, fish, hydropower and are the most important transport component (water transport accounts for about 10% of the region's cargo turnover). The most developed economic activities are metallurgy, mining and petrochemical industries, timber industry and transport. Agriculture and the service sector are poorly developed.

The main fuel in the Siberian Federal District is coal. This is due, firstly, to the development of coal mining in the region, and secondly, to the lack of access to pipeline gas. The Kovyktinskoye gas condensate field is located on the territory of Irkutsk Oblast - the largest in terms of gas reserves in eastern Russia (1.8 trillion m^3, [84]). This field, together with Chayandinsky (Yakutia), form a resource base for gas supplies to China via the «Power of Siberia» gas pipeline. The field is scheduled to be brought to full capacity in 2033.

Irkutsk Oblast was included in the gasification programme of the regions of Russia only in [85]. At the St. Petersburg International Economic Forum in June 2022, the Governor of Irkutsk Oblast, Igor Kobzev, and the Chairman of the «Gazprom» Management Board, Alexey Miller, signed an updated programme for the development of gas supply and gasification of Irkutsk Oblast for 2021-2025 [86].

2.6.1. Overview of the main potential natural gas consumers in the region

Thermal power plants and CHP. Centralized electricity production in the region is carried out by 4 hydroelectric power plants and 14 thermal power plants, while HPPs account for 69.7% of the capacity, and CHP plants account for 30.3%. The total capacity of all power plants is 13 065.8 MW. All thermal power plants in the region use steam boilers and turbine units, while 99% of the fuel used is solid fuel (mainly coal). There are also about 40 small power plants in the region that supply settlements isolated from the centralized energy system (CES). Their combined

[27] As of January 1, 2022, [83].

capacity is 19.1 MW.

When converting CHP plants from solid fuel to natural gas, it is necessary to ensure the preservation of the following three parameters of steam boilers: steam capacity, steam outlet pressure and steam outlet temperature. In this case, only the part of the CHP equipment that is responsible for steam generation is subject to modernization. It should be noted separately that it is also possible to use coal-fired boilers with minor modernization. In the latter case, the efficiency of boilers, we assume, does not change.

In 2021, the HPPs and CHP plants of Irkutsk Oblast produced 65 040 million kW·h, of which 11 650 million kW·h were generated at CHP (17.9%). During the same year, the electricity consumption in the Central Heating system of Irkutsk Oblast amounted to 59 256.2 million kW·h. In addition to electricity, CHP produced 25.25 million Gcal of thermal energy in 2021 (68.9% of the total output in the region, excluding electric boiler installations and individual heating furnaces).

About 7 829.1[28] thousand t.c.f. of coal and other solid fuels (firewood, wood chips) were consumed as fuel at the CHP in 2021, the rest (about 1%) is liquid fuel and gas. Assuming that the efficiency of the boiler equipment is maintained, 6.78 billion m^3 of gas per year will be required to completely replace coal and other solid fuels at CHP with gas. Detailed data on all thermal power plants in Irkutsk Oblast are shown in table 5.

[28] This value is approximate, since data for 2021 were not available for 2 of the 14 CHP plants, data for previous years were used for them.

Table 5. Indicators of thermal power plants of Irkutsk Oblast for 2021

Name	**Coordinates**	**Consumption of coal and other solid fuels, t.c.f.(v_s)**	**Electricity generated, million kWt·ч**	**Heat release, Gcal**	**Installed electrical power, MW**	**Installed thermal capacity GCal/h**
Novo-Ziminskaya TPP (NZTPP)	54.032, 102.032	564 022	1 086.30	1 570 277	260	818.7
Novo-Irkutsk TPP (NITPP)	52.247, 104.204	1 286 090	2 615.16	5 141 745	708	1 729.1
TI&TS Sector of the Irkutsk CHP-6 (CHP-7)	56.306, 101.725	216 998	65.35	1 356 475	12	300.8
Shelekhovsky section of Novo-Irkutskaya TPP (TPP-5)	52.182, 104.093	119 824	88.82	632 154	18	346.7
Irkutsk TPP-6	56.122, 101.602	517 992	646.92	2 291 939	270	1 442.6
Irkutsk TPP-9	52.532, 103.936	1 404 918	1 790.02	6 159 389	540	2398.6
Irkutsk TPP-10	52.461, 103.978	1 164 752	2 991.47	325 176	1 110	563.0
Irkutsk TPP-11	52.782, 103.603	340 706	543.43	948 676	320.3	1 056.9
Irkutsk TPP-12	53.124, 103.134	67 588	51.80	363 545	12	217.5
Irkutsk TPP-16	56.570, 104.141	97 072	77.23	530 208	18	322.5
Ust-Ilimskaya TPP	58.049, 102.806	438 623	779.40	1 810 627	515	1 015.0
TPP LLC «Teplosnabzhenie» (Baikal TPP)	51.515, 104.182	51 247	59.13	n/a	24	282.8
TPP of JSC «Ilim Group» in Bratsk	56.111, 101.457	919 729 (estimate)	500.80	n/a	101	n/a
TPP of JSC «Ilim Group in Ust-Ilimsk	58.043, 102.805	639 520 (estimate)	348.88	n/a	44.4	n/a
Total		**7 829 081**	**11 644.71**	**21 130 211**	**3 952.7**	**10 494.2**

Thermal boilers. In Irkutsk oblast, at the beginning of 2021, 991 boiler houses and a large number of heat recovery units (HRU) and individual heating furnaces (IHF) are involved in the production of thermal energy, in addition to thermal power plants. Of the total heat generated in 2021 (excluding electric boiler installations and IHF), amounting to 36.66 million Gcals of thermal energy, 11.41 million Gcals (31.1%) were generated by boilers and IHF.

Boiler houses of Irkutsk Oblast consumed 1 137.2 thousand t.c.f. of coal, 112.0 thousand t.c.f. of liquid fuel, 21.0 thousand t.c.f. of electricity, 7.5 thousand t.c.f. of gas and 82.0 thousand t.c.f. of other solid fuels in 2019[29]. To completely replace coal and other solid fuels in boilers with natural gas, 1.06 billion m^3 of gas per year will be required. Detailed data on boiler houses are given in table 6.

[29] At the time of the study, data on fuel consumption by boiler houses in Irkutsk Oblast for later years were not available.

Table 6. Indicators of boiler houses of Irkutsk Oblast for 2019

Name of the municipality	**Coordinates**	**The area of the municipality, km^2 (2020, S_i^b)**	**Number of boiler houses, pcs. (2019)**	**Coal consumption, tons (2019)**	**Consumption of other olid fuels, (tons, 2019)**	**Consumption of coal and other solid fuels, t.c.f. (assessment, 2019, v_i^b)**
Bratsk	56.152, 101.634	432	7	96 508	1 499	59 063
Zima	53.921, 102.049	53	11	31 024	0	18 859
Irkutsk	52.290, 104.281	277	34	401 378	0	243 987
Svirsk	53.067, 103.342	39	4	45 474	0	27 642
Tulun	54.557, 100.578	134	24	108 399	0	65 893
Ust-Ilimsk	57.959, 102.735	227	3	0	0	0
Cheremkhovo	53.137, 103.090	119	19	12 170	0	7 398
Angarsk district	52.349, 103.695	1 149	3	0	0	0
Balagansky district	54.668, 102.888	6 347	21	3 899	1 927	2 883
Bodaibinsky district	58.522, 115.853	92 171	22	77 418	0	47 060
Bratsky district	55.720, 101.761	30 838	56	96 127	12 985	61 888
Zhigalovsky district	54.945, 105.364	21 818	14	3 114	2 774	2 631
Zalarinsky district	53.281, 101.570	7 617	30	35 659	0	21 676
Ziminsky district	53.760, 101.539	7 019	27	11 427	0	6 946
Irkutsk district	52.172, 104.897	11 688	36	33 221	1 251	20 527
Kazachinsko-Lensky district	55.781, 107.836	33 276	11	40 069	2 497	25 021
Katanga district	60.832, 107.434	139 163	8	0	0	0
Kachugsky district	54.109, 106.627	31 395	42	3 837	20 326	7 741
Kirensky district	57.978, 109.212	43 905	18	10 285	52 940	20 339
Kuytunsky district	54.529, 101.600	11 147	44	11 795	2 608	7 864
Mamsko-Chuisky district	58.106, 112.291	43 134	9	27 027	0	16 429
Nizhneilimsky district	56.865, 103.759	18 879	26	16 651	105 205	38 115
Nizhneudinsky district	54.589, 99.106	49 564	79	92 715	11 028	59 294
Olkhonsky district	53.091, 107.194	15 895	13	5 500	1 200	3 663
Slyudyansky district	51.657, 103.718	6 301	21	149 802	0	91 061
Taishet district	55.553, 97.814	27 725	63	203 194	0	123 516
Tulunsky district	54.178, 100.275	13 870	38	12 421	1 060	7 833
Usolsky district	52.416, 103.215	6 252	37	33 099	244	20 185
Ust-Ilimsky district	57.593, 101.640	36 596	13	23 753	20 485	19 889
Ust-Kutsky district	56.950, 105.911	34 604	21	152 743	57 173	108 061
Ust-Udinsky district	55.135, 103.890	20 110	18	6 435	3 550	4 856
Cheremkhovsky district	52.772, 102.282	9 926	23	40 995	1 040	25 196
Chunsky district	56.765, 99.702	28 036	31	47 845	8 377	31 313
Shelekhovsky district	52.029, 103.848	2 014	16	6 464	0	3 929
Alarsky district	53.318, 102.759	2 651	36	9 921	0	6 030
Bayandaevsky district	53.138, 105.571	3 756	20	3 358	0	2 041
Bokhansky district	53.151, 103.920	3 678	32	4 900	0	2 979
Nukutsky district	53.751, 102.881	2 473	22	4 082	0	2 481
Osinsky district	53.603, 104.042	4 402	18	1 300	0	790
Ehirit-Bulagatsky district	52.922, 104.919	5 153	25	6 777	0	4 120
Total			**995**	**1 870 785**	**308 169**	**1 219 200**

The rest of the consumption. In 2020[30], the consumption of coal and other solid fuels in Irkutsk Oblast, excluding consumption by thermal power plants and boiler houses, amounted to only 603[31] thousand. t.c.f. At the same time, there is no detailed consumption in the available data for various settlements and enterprises. Therefore, when estimating the demand function for natural gas for Irkutsk Oblast, it was decided not to take into account potential consumption by enterprises and rural populations.

2.6.2. Preparation of the initial task parameters

To assess the prospects for gasification of Irkutsk Oblast, we will build a model of the natural gas market consisting of producers, consumers and the transport system. We assume that the expected interest rates on bank deposits are equal to the levels of inflation. The calculation is carried out in prices at the beginning of 2022.

As production nodes, we use two gas fields that are the resource base for the «Power of Siberia» gas pipeline: Kovyktinskoye and Chayandinskoye (table 7). Let's estimate the function of production costs by the ratio (2.11), i.e. we consider that the marginal production costs are constant and equal to 2 095 rub./t.c.f.

Table 7. Producing nodes of the natural gas market of Irkutsk Oblast

Node number	**Producer**	**Coordinates (approximate location)**
1	Kovyktinskoye field	55.385, 106.124
2	Chayandinskoye field	59.826, 110.919

We will consider 14 thermal power plants and 40 municipalities of the region as consumers. For the latter, we will take into account the demand from boiler houses. We will set the demand functions for natural gas according to formulas (2.15) and (2.16, 2.17), using the estimate obtained in paragraph 2.3.2. for marginal transport costs for gas pipelines of distribution networks $e^d =$ 485.8 (rub./(t.c.f.×km)) and information from tables 5 and 6. To estimate the cost of extraction and delivery of a fuel unit used (parameters c_s and c_i^b), we take into account data from [87], according to which in 2011 the cost of coal supply for the CHPP of Irkutsk Oblast was 347-1 042 rub. per ton, while prices differed for various types of coal and different deposits. In prices at the

[30] At the time of the study, data for later years were not available.
[31] This value is approximate, since there are significant statistical discrepancies in the data on resource consumption in [2].

beginning of 2022, the cost of coal supply would be 1 140-3 425 rub./t.c.f. Thus, the average cost[32] of extracting and delivering a unit of coal, excluding the producer's margin, can be taken to be 1 500-2 500 rub./t.c.f.

It is also worth considering the environmental component associated with the fact that the level of environmental pollution when using coal significantly exceeds similar indicators for natural gas, which is considered an environmentally friendly type of fuel. Let's take this component into account by adding a penalty modifier to the parameters c_s and c_i^b. Let's consider 7 options for the values of these parameters: 1 500, 2 000, 2 500, 3 000, 3 500, 5 000, 7 000 (rub./t.c.f.).

To estimate the penalty modifier, you can use the amount of the carbon tax used in many countries. This tax is levied when carbon dioxide (CO_2) is released into the atmosphere and is proportional to the volume of emissions. The rate of this tax can vary greatly for different countries. For example, as of April 1, 2022, the tax rate in France was 45 euros per ton of CO_2 [88]. It is known that to generate the same amount of thermal energy, coal combustion emits more CO_2 into the atmosphere than natural gas combustion. According to [63], this difference is 1 171 kg. CO_2 per t.c.f. of generated heat. This means that with a tax rate of 45 euros per ton of CO_2, this difference corresponds to a tax of 52.7 euros or 4486 rubles (according to the exchange rate as of January 1, 2022), which, when taking into account the costs of mining and shipping a unit of coal, corresponds to scenarios 6-7. However, it must be borne in mind that burning coal also leads to the release of other pollutants into the atmosphere, in addition to carbon dioxide. These emissions can also be taken into account in the penalty modifier.

We will divide all consuming nodes into three categories according to the total fuel consumption:

- consumption of over 100 000 t.c.f./year (category 1);
- consumption in the range of 20 000-100 000 t.c.f./year (category 2);
- consumption is less than 20 000 tons per year (category 3).

Data on consuming nodes are presented in tables 8 and 9.

[32] When estimating average costs, the influence of the remoteness of the consuming node from the coal mining site was not taken into account, thus the average cost of purchasing tons of coal was assumed to be the same for all nodes.

Table 8. Consuming nodes of the Irkutsk Oblast natural gas market (TTP)

Node number	**Consumer**	**Coordinates**	v_s **(fuel consumption per plant, t.c.f./year)**	**Category**
3	Novo-Ziminskaya TPP (NZTPP)	54.032, 102.032	564 022	1
4	Novo-Irkutskaya TPP (NITPP)	52.247, 104.204	1 286 090	1
5	TI&TS of the Irkutsk TPP-6 (TPP-7)	56.306, 101.725	216 998	1
6	Shelekhovsky section of Novo-Irkutskaya TPP (TPP-5)	52.182, 104.093	119 824	1
7	Irkutsk TPP-6	56.122, 101.602	517 992	1
8	Irkutsk TPP-9	52.532, 103.936	1 404 918	1
9	Irkutsk TPP-10	52.461, 103.978	1 164 752	1
10	Irkutsk TPP-11	52.782, 103.603	340 706	1
11	Irkutsk TPP-12	53.124, 103.134	67 588	2
12	Irkutsk TPP-16	56.570, 104.141	97 072	2
13	Ust-Ilimskaya TPP	58.049, 102.806	438 623	1
14	TPP LLC «Teplosnabzhenie» (Baikal TPP)	51.515, 104.182	51 247	2
15	TPP of JSC «Ilim Group» in Bratsk	56.111, 101.457	919 729 (estimate)	1
16	TPP of JSC «Ilim Group in Ust-Ilimsk	58.043, 102.805	639 520 (estimate)	1
Total			**7 829 081**	

Table 9. Consuming nodes of the natural gas market of Irkutsk oblast (boiler houses)

Node number	**Consumer**	**Coordinates**	$e^d\sqrt{\frac{S_i^b}{\pi}}$ $\left(\frac{\textbf{rub.}}{\textbf{t.c.f.}}\right)$	$\frac{\pi \cdot v_i^b}{S_i^b \cdot e^{d2}}$ $\left(\frac{\textbf{t.c.f.}^3}{\textbf{rub.}^2 \times \textbf{year}}\right)$	v_i^b **(total fuel consumption of the boiler houses, t.c.f./year)**	**Category**
17	Bratsk	56.152, 101.634	5 699	1.818×10^{-3}	59 063	2
18	Zima	53.921, 102.049	1 993	4.750×10^{-3}	18 859	3
19	Irkutsk	52.290, 104.281	4 565	1.171×10^{-2}	243 987	1
20	Svirsk	53.067, 103.342	1 703	9.528×10^{-3}	27 642	2
21	Tulun	54.557, 100.578	3 167	6.569×10^{-3}	65 893	2
22	Ust-Ilimsk	57.959, 102.735	4 127	0	0	3
23	Cheremkhovo	53.137, 103.090	2 985	8.303×10^{-4}	7 398	3
24	Angarsk district	52.349, 103.695	9 289	0	0	3
25	Balagansky district	54.668, 102.888	21 836	6.046×10^{-6}	2 883	3
26	Bodaibinsky district	58.522, 115.853	83 211	6.797×10^{-6}	47 060	2
27	Bratsky district	55.720, 101.761	48 131	2.672×10^{-5}	61 888	2
28	Zhigalovsky district	54.945, 105.364	40 484	1.605×10^{-6}	2 631	3
29	Zalarinsky district	53.281, 101.570	23 921	3.788×10^{-5}	21 676	2
30	Ziminsky district	53.760, 101.539	22 963	1.317×10^{-5}	6 946	3
31	Irkutsk district	52.172, 104.897	29 632	2.338×10^{-5}	20 527	2
32	Kazachinsko-Lensky district	55.781, 107.836	49 997	1.001×10^{-5}	25 021	2
33	Katanga district	60.832, 107.434	102 246	0	0	3
34	Kachugsky district	54.109, 106.627	48 564	3.282×10^{-6}	7 741	3
35	Kirensky district	57.978, 109.212	57 430	6.167×10^{-6}	20 339	2
36	Kuytunsky district	54.529, 101.600	28 938	9.391×10^{-6}	7 864	3
37	Mamsko-Chuisky district	58.106, 112.291	56 924	5.070×10^{-6}	16 429	3
38	Nizhneilimsky district	56.865, 103.759	37 659	2.688×10^{-5}	38 115	2
39	Nizhneudinsky district	54.589, 99.106	61 019	1.592×10^{-5}	59 294	2
40	Olkhonsky district	53.091, 107.194	34 555	3.067×10^{-6}	3 663	3
41	Slyudyansky district	51.657, 103.718	21 757	1.924×10^{-4}	91 061	2
42	Taishet district	55.553, 97.814	45 637	5.930×10^{-5}	123 516	1
43	Tulunsky district	54.178, 100.275	32 279	7.517×10^{-6}	7 833	3
44	Usolsky district	52.416, 103.215	21 671	4.298×10^{-5}	20 185	2
45	Ust-Ilimsky district	57.593, 101.640	52 432	7.235×10^{-6}	19 889	3
46	Ust-Kutsky district	56.950, 105.911	50 986	4.157×10^{-5}	108 061	1
47	Ust-Udinsky district	55.135, 103.890	38 868	3.215×10^{-6}	4 856	3
48	Cheremkhovsky district	52.772, 102.282	27 306	3.379×10^{-5}	25 196	2
49	Chunsky district	56.765, 99.702	45 892	1.487×10^{-5}	31 313	2
50	Shelekhovsky district	52.029, 103.848	12 299	2.597×10^{-5}	3 929	3
51	Alarsky district	53.318, 102.759	14 112	3.028×10^{-5}	6 030	3
52	Bayandaevsky district	53.138, 105.571	16 798	7.234×10^{-6}	2 041	3
53	Bokhansky district	53.151, 103.920	16 622	1.078×10^{-5}	2 979	3
54	Nukutsky district	53.751, 102.881	13 631	1.336×10^{-5}	2 481	3
55	Osinsky district	53.603, 104.042	18 185	2.390×10^{-6}	790	3
56	Ehirit-Bulagatsky district	52.922, 104.919	19 675	1.064×10^{-5}	4 120	3
Total					**1 219 200**	

We will build a transport system for the market using the section of the «Power of Siberia» gas pipeline passing through the Kovyktinskoye and Chayandinskoye fields. We assume that this section has a sufficient reserve capacity, and the cost of transporting 1 t.c.f. of natural gas per 1 km along it is 0.739 rub. (see paragraph 2.3.1.). Let's combine all the producing and consuming nodes into a single tree network in three stages. First, we connect the consuming nodes of the first category with the producing nodes, while sequentially connecting the nearest isolated consuming node to the network. After that, we will perform a similar procedure for consuming nodes of the second and third categories. The resulting scheme of the transport system consists of 6 existing and 76 potential lines and is shown in figures 35-40. Nodes numbered 57-83 are intermediate.

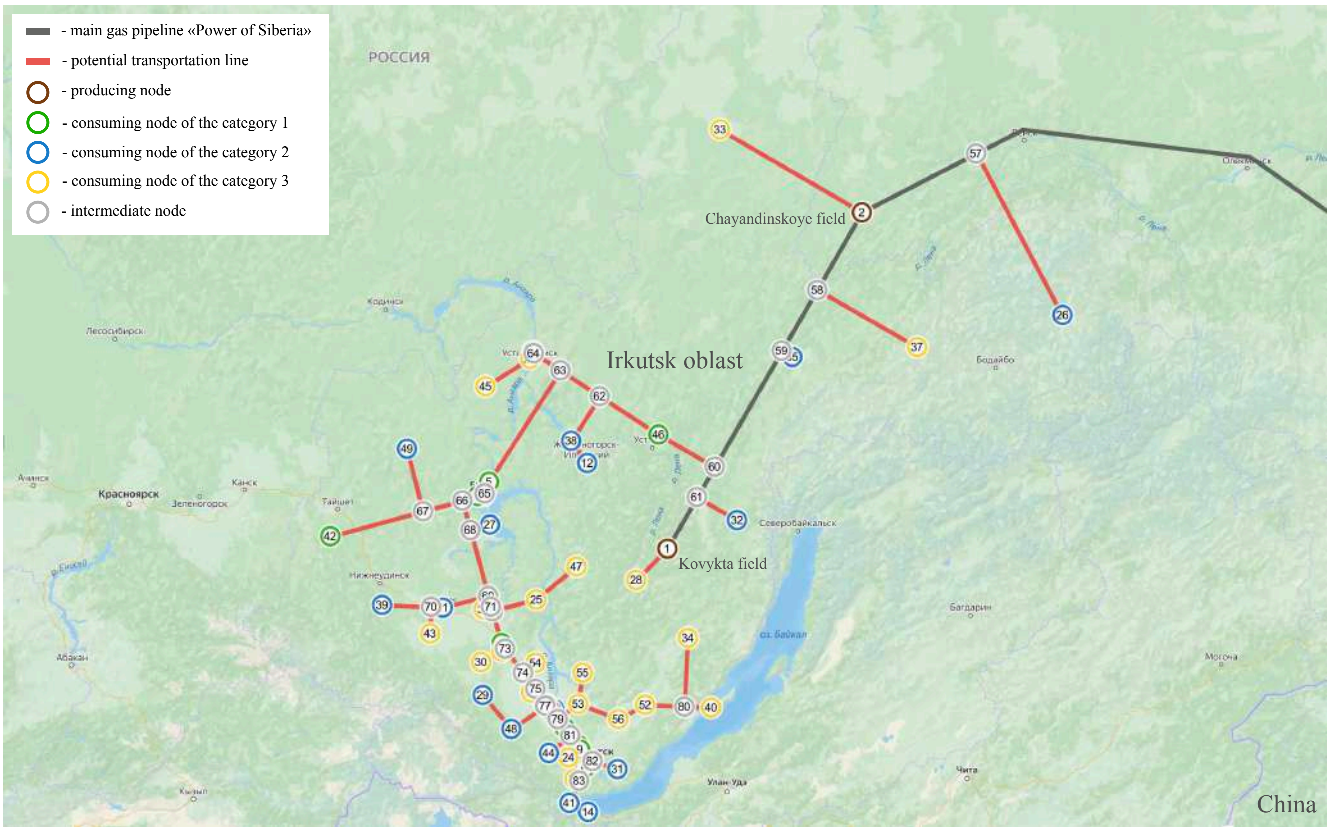

Fig. 35. The scheme of the transport system of the natural gas market of Irkutsk Oblast (complete)

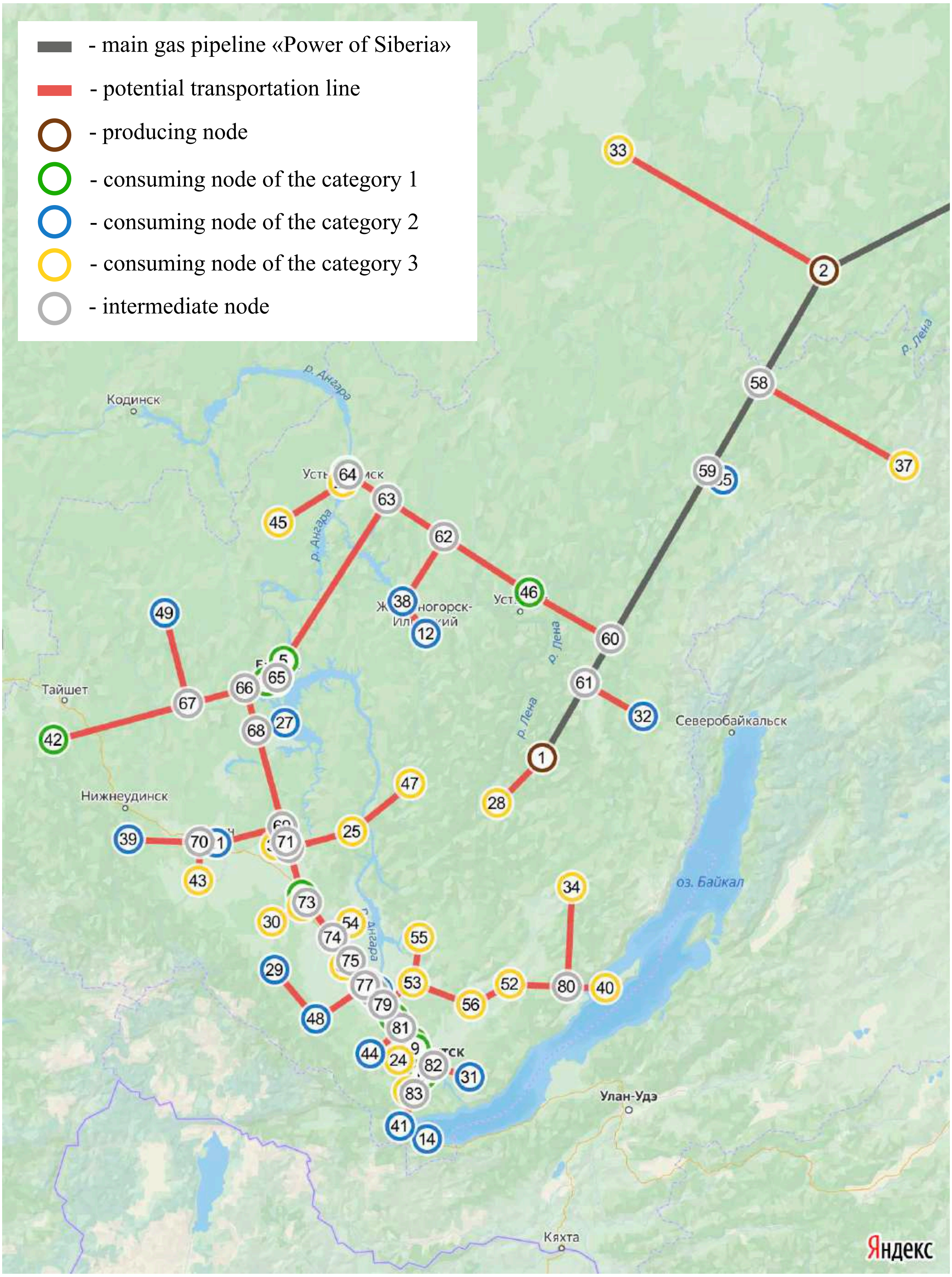

Fig. 36. The scheme of the transport system of the natural gas market of Irkutsk Oblast (enlarged)

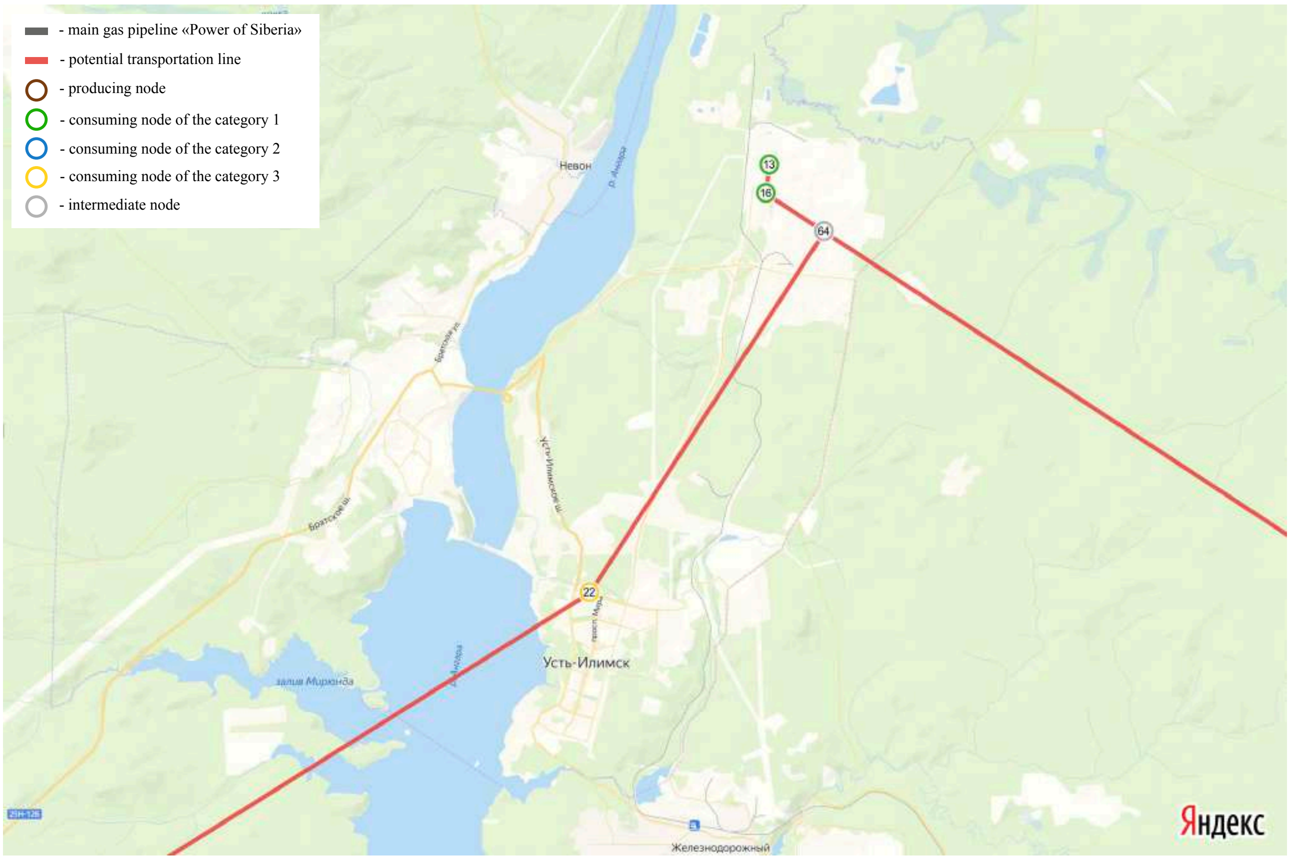

Fig. 37. The scheme of the transport system of the natural gas market of Irkutsk Oblast (Ust-Ilimsk)

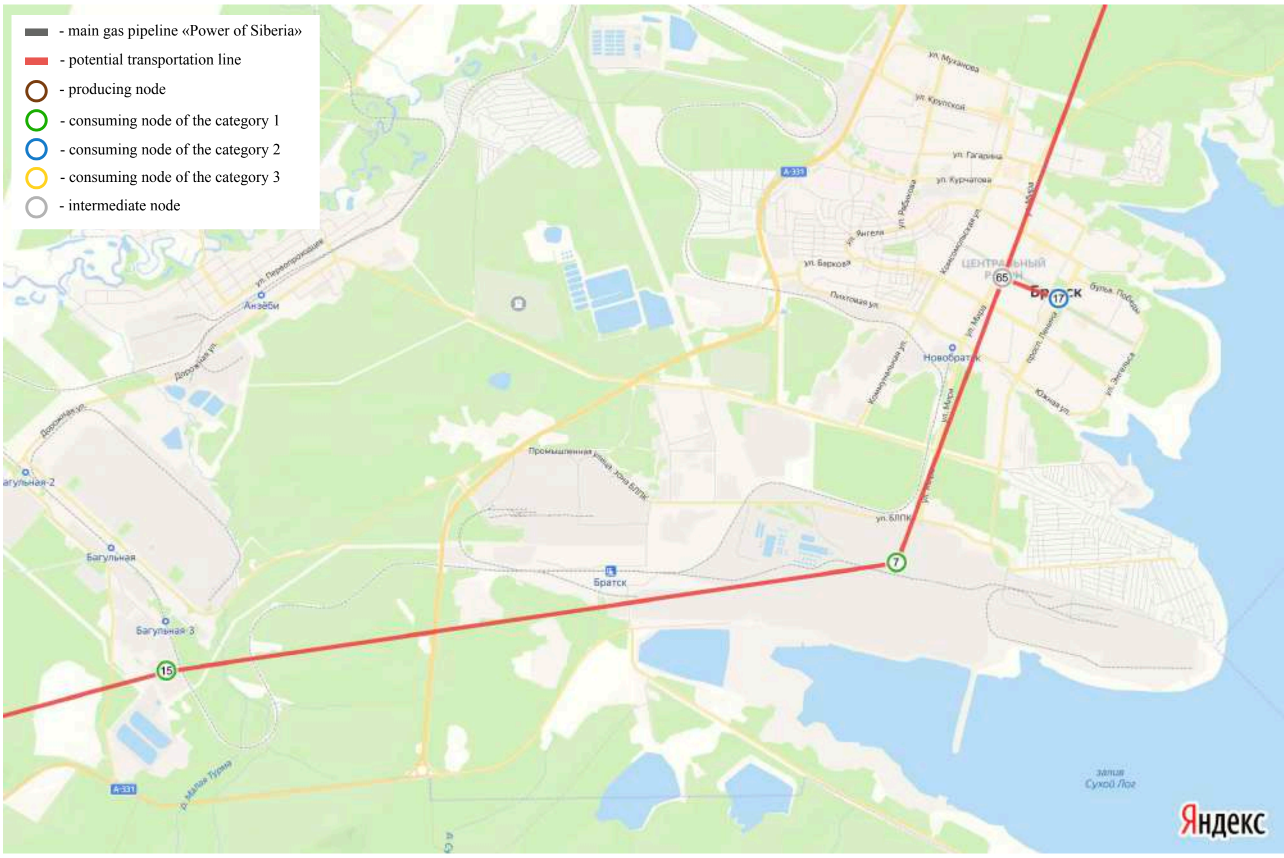

Fig. 38. The scheme of the transport system of the natural gas market of Irkutsk Oblast (Bratsk)

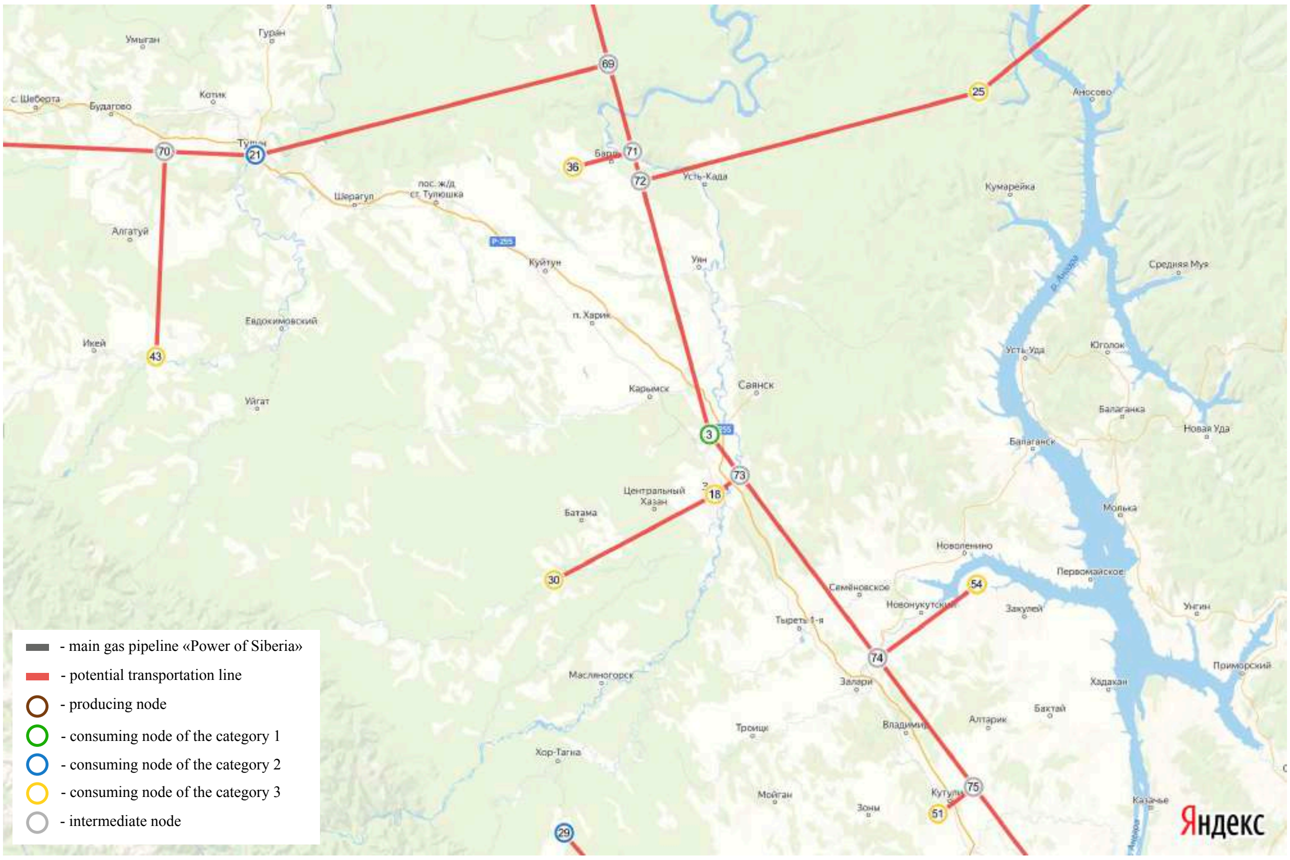

Fig. 39. The scheme of the transport system of the natural gas market of Irkutsk Oblast (Tulun)

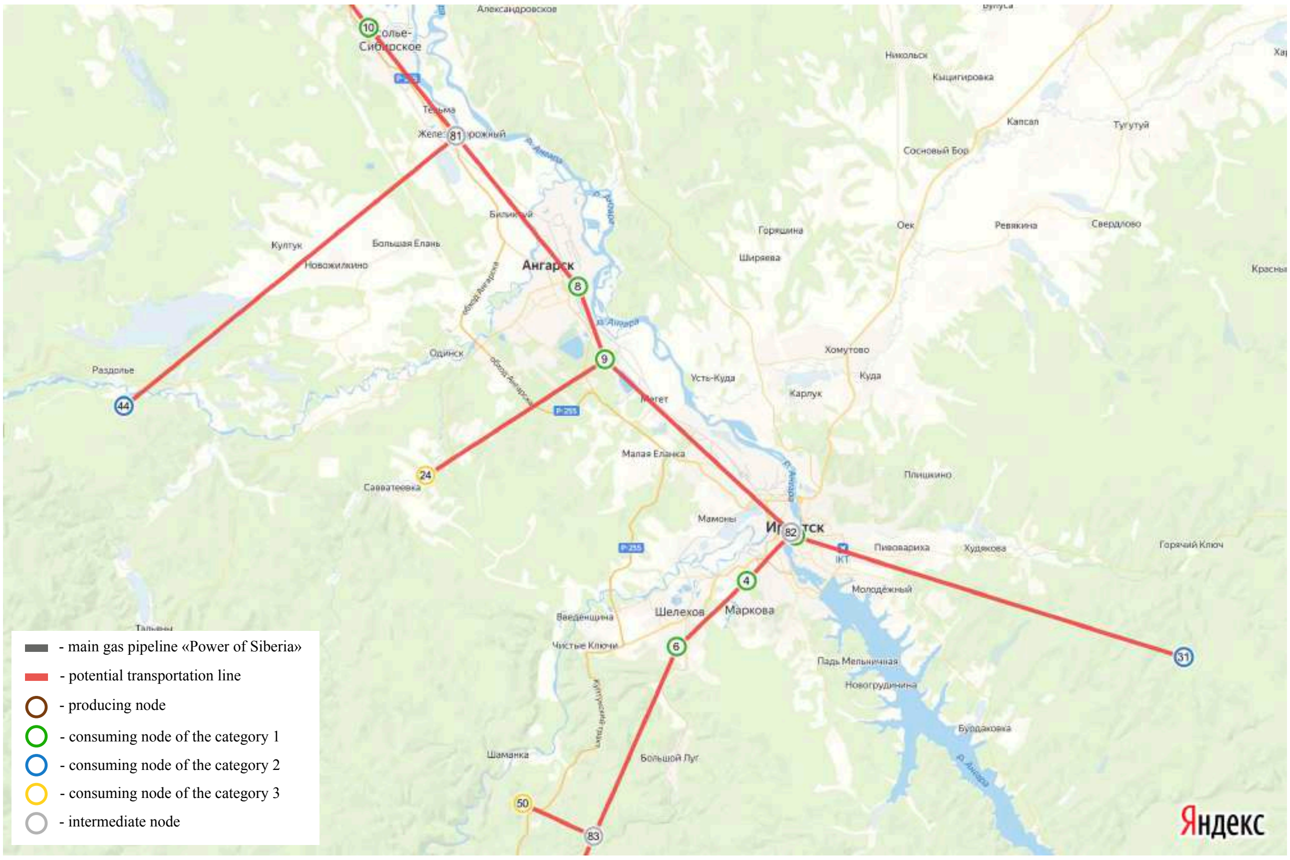

Fig. 40. The scheme of the transport system of the natural gas market of Irkutsk Oblast (Irkutsk)

As functions of transport costs for potential lines, we use the estimate (2.2), while assuming that the length of each line is 25% longer than the shortest distance between its incident nodes. Detailed descriptions of the lines are provided in the Appendices (tables 11, 12).

2.6.3. Calculation results

This section describes the results of applying the developed algorithms described in the first chapter of the work to optimize the constructed model of the natural gas energy market in Irkutsk Oblast. For 7 scenarios with different values of the cost of extraction and delivery of a fuel unit used, taking into account the environmental component (parameters c_s and c_i^b), the main calculation results are presented in table 10. Detailed results are presented in the Appendices (tables 13-15).

Table 10. Main results of optimization of the transport system of the natural gas market of Irkutsk Oblast for various scenarios

<table>
<tr><th></th><th>Scenario 1</th><th>Scenario 2</th><th>Scenario 3</th><th>Scenario 4</th><th>Scenario 5</th><th>Scenario 6</th><th>Scenario 7</th></tr>
<tr><td>Cost of extraction and delivery of a fuel unit used, taking into account the environmental component, rub./t.c.f.</td><td>1 500</td><td>2 000</td><td>2 500</td><td>3 000</td><td>3 500</td><td>5 000</td><td>7 000</td></tr>
<tr><td>Economic effect, mln rub./year</td><td colspan="4">0</td><td>656.4</td><td>11 841.7</td><td>27 599.3</td></tr>
<tr><td>Number of expandable lines</td><td colspan="4">0</td><td>10</td><td>33</td><td>38</td></tr>
<tr><td>Total length of expandable lines, km</td><td colspan="4">0</td><td>686.5</td><td>1 348.5</td><td>1 511.4</td></tr>
<tr><td>Total flow on all lines, mln (t.c.f.×km)/year</td><td colspan="4">0</td><td>1 944.7</td><td>8 829.5</td><td>9 115.8</td></tr>
<tr><td>Number of nodes with positive consumption</td><td colspan="4">0</td><td>6</td><td>16</td><td>20</td></tr>
<tr><td>Total consumption of natural gas, thousand t.c.f. per year</td><td colspan="4">0</td><td>2 732.9</td><td>7 750.4</td><td>8 023.5</td></tr>
<tr><td>The average number of auxiliary problems solved</td><td colspan="3">305</td><td>1 169</td><td>71 285</td><td>27 667</td><td>15 659</td></tr>
</table>

As can be seen from the table, with low costs for the extraction and delivery of a unit of the type of fuel used (scenarios 1-4), gasification of Irkutsk Oblast is impractical. At costs from 3 500 rub./t.c.f. gasification brings a positive annual effect, which is equal to 656.4 mln rub., 11.8 billion rub. and 27.6 billion rub. for scenarios 5, 6 and 7, respectively. However, these three scenarios are implemented only if the environmental component is taken into account.

Figures 41, 42 show a diagram of a transport system consisting only of lines the expansion of which is optimal for scenario 7. For this scenario, the optimal plan for the development of the transport system involves connecting all thermal power plants with the Kovyktinskoye field, except for Baikal, the cities of Bratsk, Zima, Irkutsk, Svirsk, Cheremkhovo, as well as Nizhneilimsky and Ust-Kutsky districts. At the same time, the CHP is undergoing a complete transition to natural gas, and cities and districts are partially gasified.

The estimates obtained are rather rough due to the numerous simplifications and assumptions that had to be made, including due to the lack of access to up-to-date detailed data. Therefore, the results of the analysis of the prospects for gasification of Irkutsk oblast should not be taken as a kind of recommendation, rather as an example of using the developed algorithms and methods.

It is also worth paying attention to the number of auxiliary tasks to be solved, which directly affects the time to solve the original problem[33]. This number is the largest for scenario 5 (71 285 $\approx 2^{16}$). In the case of a complete search of all possible sets of expandable lines, this number would be $2^{76} \approx 10^{23}$ (76 is the number of potential transmission lines). Thus, the calculations performed showed that the developed algorithms can be used to plan the development of real energy markets with a large number of nodes and find the optimal plan for the development of the transport system, maximizing social welfare, in a reasonable time[34].

[33] As it was proved in paragraph 1.3.7., when certain conditions are met, the complexity of the algorithm for solving the auxiliary problem is proportional to the square of the number of nodes in the market.

[34] Scenario 5 was calculated in 6.5 seconds on the personal computer of the study's author.

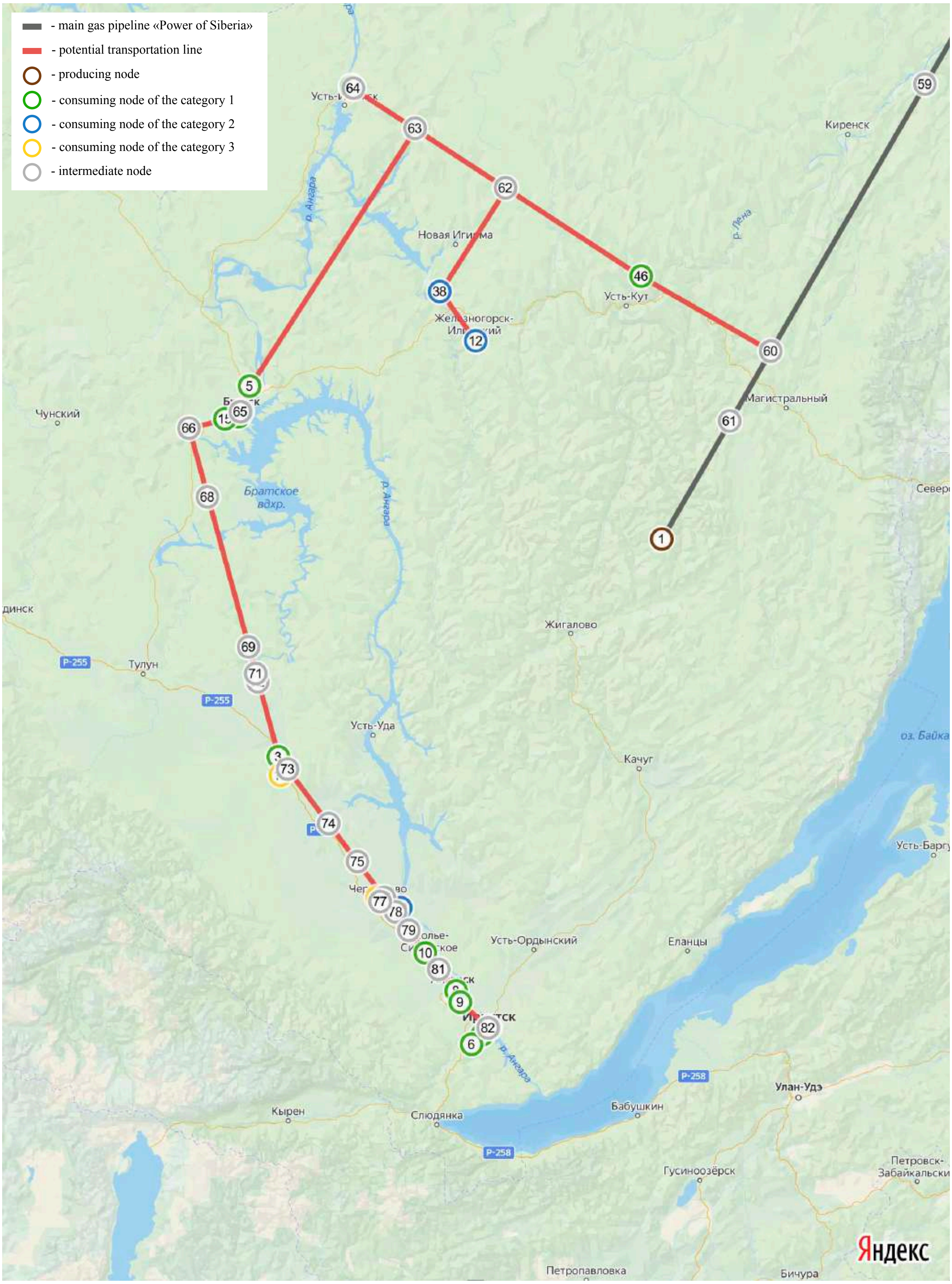

Fig. 41. Optimal set of expandable lines for Irkutsk Oblast natural gas market for scenario 7 (complete; non-expandable lines were removed from the original scheme)

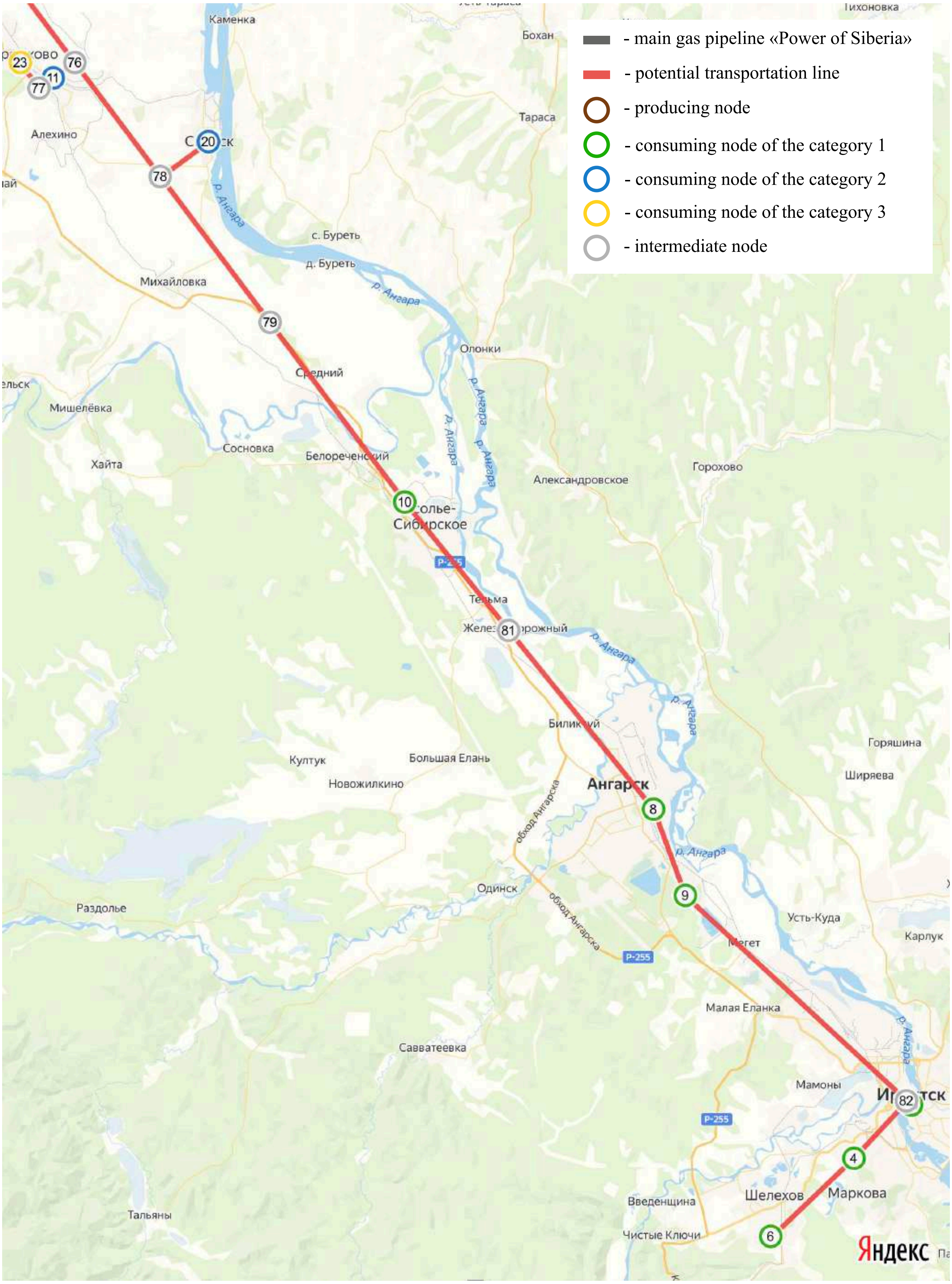

Fig. 42. The optimal set of expandable lines for Irkutsk Oblast natural gas market for scenario 7 (Irkutsk; non-expandable lines were removed from the original scheme)

Conclusions to the second chapter

In this chapter, methods for estimating the initial parameters for the energy market model studied in the work were proposed. These parameters are the functions of production costs, demand and transmission costs. Although these methods have been developed for the natural gas market, they can also be adapted for other markets.

Two types of gas pipelines are considered: main and distribution pipelines. For each type, an estimate of the transport cost function for the construction of a new gas pipeline is derived. This estimate depends on the capacity and length of the gas pipeline.

For a gas field, the problem of modelling the production cost function is investigated. It is proposed to calculate production costs according to one of two models. The first model was actively used to perform calculations for many groups of gas fields in the USSR. The second model is simplified, it assumes that marginal production costs are constant. The corresponding estimate is derived for the latter.

To assess the demand function in a node that is an arbitrary non-gasified entity or municipality of the Russian Federation, a method has been developed based on dividing the main gas consumers into several groups and evaluating the corresponding component of the demand function for each group.

The developed algorithms and methods were applied to assess the prospects of gasification of Irkutsk Oblast. The possibility of connecting thermal power plants and boiler houses in the region to the main gas pipeline «Power of Siberia» was considered. According to the calculations based on data on the characteristics of thermal power plants and boiler houses in the region for 2021-2022, gasification brings a positive effect only if the environmental component is taken into account, expressed in the form of a fine for burning each unit of coal currently used in the region. The calculations also showed that the algorithms developed in the first chapter of the work enable solving the initial problem of maximizing social welfare in a reasonable time, even for a large number of nodes, which enables using these algorithms to plan the development of real energy markets.

Conclusion

In the course of the study, the following main scientific results were obtained:

1) • for a multi-node energy market of a single resource of the star-type (and the more general case of the tree-type market), in which fixed costs are present when expanding transport lines that do not depend on the volume of expansion, the initial task of optimizing the transport system in terms of maximizing social welfare is NP-hard;

2) • for the auxiliary task of optimizing a transport system with a fixed set of expandable lines, which is a convex programming problem, a special solution algorithm has been developed;

• its complexity is estimated for the case of piecewise linear initial functions: the number of computational operations of the algorithm does not exceed the value of some predetermined quadratic function of the number of nodes in the market;

3) • algorithms for solving the initial problem for various transport structures have been developed for the case of the flow structure invariance condition being met, in which the directions of flows in transmission lines are constant and do not depend on transmission capacities; algorithms have been developed for the following transport structures: «chain», «star», «star-chain»;

• the average statistical complexity of these algorithms is investigated: for computational experiments with random generation of initial problems for each of the three cases, even for a large number of nodes (more than 50), the initial problem is solved in a reasonable time, and the dependence of the average number of solved auxiliary problems on the number of nodes in the market is approximated by a quadratic function;

4) • for the natural gas market, methods have been developed for estimating the demand functions for non-gasified nodes;

5) • the developed algorithms and methods are applied to assess the prospects of gasification of Irkutsk oblast with the possibility of connecting thermal power plants and boiler houses of the region to the main gas pipeline «Power of Siberia»;

• according to the calculations based on data on the characteristics of thermal power plants and boiler houses in the region for 2021-2022, gasification brings a positive effect only if the environmental component is taken into account, expressed in the form of a fine for burning each unit of coal currently used in the region;

- the calculations performed showed that the developed algorithms can be used in planning the development of real energy markets and allow solving the initial problem in a reasonable time.

Thus, in the course of the study, methods of effective planning for the development of transport systems of energy markets were developed, and also, using the example of Irkutsk Oblast, it is shown how these methods can be used to assess the prospects for gasification of specific regions of the Russian Federation. The results obtained correspond to the purpose of the study and can be used in planning the development of real gas or oil transportation energy systems.

List of acronyms and symbols

- *AGPPS* - Advanced Gas Production Planning System;
- *AS USSR* - Academy of Sciences of the USSR;
- *BCS* - booster compressor station;
- *CEMI RAS* - Central Economic and Mathematical Institute of the Russian Academy of Sciences;
- *CHP* - combined heat and power;
- *etc.* - so on;
- *FAS* - Federal Antimonopoly Service;
- *FSIC* - the flow structure invariance condition;
- *Gcal* - gigacalory;
- *GPP* - gas processing plant;
- *HAC* - Higher Attestation Commission;
- *HPS* - hydroelectric power station;
- *HRU* - heat recovery unit;
- *IEEE* - Institute of Electrical and Electronics Engineers;
- *IHF* - individual heating furnace;
- *IMP RAS* - Institute of Management Problems of the Russian Academy of Sciences;
- *J* - joule;
- *JSC* - joint-stock company;
- *kcal* - kilocalory;
- *kg* - kilogram;
- *km* - kilometer;
- km^2 - square kilometer;
- *kW* - kilowatt;
- *LLC* - limited liability company;
- m^3 - cubic meter;
- *MJ* - megajoule;
- *MRI* - magnetic resonance imaging;
- *MW* - megawatt;
- *n.* - number;
- *NITPP* - Novo-Irkutskaya TPP;
- *NP* - non-deterministic polynomial;
- *NPP* - nuclear power plant;
- *NZTPP* - Novo-Ziminskaya TPP;
- *p.* - page;
- *REC* - Russian Economic Congress;
- *RES* - renewable energy sources;
- *RF* - Russian Federation;
- *rub.* - ruble;
- *t* - ton;
- *t.c.f.* - ton of conventional fuel;
- *TPP* - thermal power plant;
- *USA* - United States of America;
- *USSR* - Union of Soviet Socialist Republics;
- *vol.* - volume;
- *№* - number.

Reference list

Regulatory acts

[1] Code of Rules (CR) dated 18.02.2014 N 86.13330.2014 «Main pipelines (revision of the updated SN&R III-42-80* «Main pipelines» (CR 86.13330.2012))» with amendments, approved by order of the Ministry of Construction of the Russian Federation of December 14, 2017. N 1664/pr «On approval of Amendment No. 2 to CR 86.13330.2014 «SN&P III-42-80* Main pipelines»».

[2] Decree of the Governor of Irkutsk Oblast of 28.04.2022 No. 71-ug «On approval of the scheme and programme for the development of the electric power industry of Irkutsk Oblast for 2023-2027» - [Digital resource]. URL: http://publication.pravo.gov.ru/Document/View/3800202204290009 (date of reference 01.05.2022).

[3] Decree of the Governor of Irkutsk Oblast of 29.042021 No. 128-ug «On approval of the scheme and programme for the development of the electric power industry of Irkutsk Oblast for 2022-2026» - [Digital resource]. URL: http://publication.pravo.gov.ru/Document/View/3800202105110005 (date of reference 01.05.2022).

[4] Decree of the Governor of Irkutsk Oblast of 29.04.2020 No. 124-ug «On approval of the scheme and programme for the development of the electric power industry of the Irkutsk Region for 2021-2025» - [Digital resource]. URL: http://publication.pravo.gov.ru/Document/View/3800202004300007 (date of reference 01.05.2022).

Scientific publications

[5] Israel, Yu. A. Chernobyl: radioactive pollution of natural environments: monograph / Yu. A. Israel, S. M. Vakulovsky, V. A. Vetrov et al. Leningrad: Gidrometeoizdat, 1990. 223 p.

[6] The Fukushima Daiichi Accident: report by the Director General of International Atomic Energy Agency. Vienna, Austria. 2015. 222 p.

[7] Arrow, K. J. Existence of an Equilibrium for a Competitive Economy / K. J. Arrow, G. Debreu // N.Y.: Econometrica. 1954. Vol. 22. P. 265–290.

[8] Walras, L. Les éléments d'économie politique pure ou Théorie de la richesse sociale. Lausanne, 1874.

[9] Wald, A. Uber einige Gleichungssysteme der mathematischen Okonomie // Zeitschrift fur Nationalokonomie. Vol. 7. 1936. P. 637–670, translated as «On Some Systems of Equations of Mathematical Economics» // N.Y.: Econometrica. Vol. 19. October, 1951. P. 368–403.

[10] Debreu, G. Valuation Equilibrium and Pareto Optimum // Proceedings of the National Academy of Sciences of the USA. 1954. Vol. 40. N. 7. P. 588–592.

[11] Crew, M. A. The Theory of Peak-Load Pricing: A Survey / M. A. Crew, C. Fernando, P. R. Kleindorfer // Journal of Regulatory Economics. 1995. Vol. 8. P. 215–248.

[12] Davidson, M. R. Mathematical model of the competitive wholesale electricity market in Russia / M. R. Davidson, Yu. V. Dogadushkina, E. M. Kreines et al. // M.: Bulletin of the Academy of Sciences (Izvestiya Akademii Nauk). Theory and control systems. 2004. № 3. P. 72–83.

[13] A mathematical model of energy system management in the conditions of a competitive wholesale electricity and capacity market in Russia / M. R. Davidson [et al.] // M.: Izvestiya RAS. Theory and control systems. 2009. Vol. 2. P. 84–94.

[14] Davidson, M. R. Mathematical Model of Power System Management in Conditions of a Competitive Wholesale Electric Power (Capacity) Market in Russia / M. R. Davidson, Y. V. Dogadushkina, E. M. Kreines, N. M. Novikova, A. V. Seleznev, Y. A. Udaltsov, L. V. Shiryaeva // M.: Journal of Computer and Systems Sciences International. 2009. N. 48. P. 243–253.

[15] Hogan, W. Competitive electricity market design: a wholesale primer: tech. rep. / W. Hogan. Massachusetts: Harvard Electricity Policy Group. 1998.

[16] Vasin, A. A. Electricity Markets Analysis and Design / A. A. Vasin, P. A. Vasina. Working Paper 2006/053: tech. rep. / Moscow New Economic School. 2006.

[17] Vogelsang, I. Price Regulation for Independent Transmission Companies / I. Vogelsang // Journal of Regulatory Economics. 2001. Vol. 20. N. 2. P. 141–165.

[18] Edoli, E. Optimization Methods for Gas and Power Markets: Theory and Cases / E. Edoli, S. Fiorenzani, T. Vargiolu. Palgrave Macmillan, Basingstoke. 2016. XVII. 192.

[19] Wu, F. Adaptive convex relaxations for gas pipeline network optimization / F. Wu, N. Harsha, A. Zlotnik, R. Sioshansi, A. M. Rudkevich // IEEE Conference Proceedings. 2017. Vol. 2017. P. 4710–4716.

[20] Roger, Z. R.-M. Optimization problems in natural gas transportation systems: a state-of-the-art review /Z. R.-M. Roger, B.-S. Conrado // Appl. Energy, 147(1). 2015. P. 536–555.

[21] Rosellon, J. Different Approaches Towards Electricity Transmission Expansion / J. Rosellon // Review of Network Economics. 2003. Vol. 2. №. 3. P. 238–269.

[22] Joskow, P. L. Transmission Rights and Market Power on Electric Power Networks / P. L. Joskow, J. Tirole // RAND Journal of Economics. 2000. Vol. 31. P. 450–487.

[23] Levit, B. Yu. Nonlinear network transport problems / B. Yu. Leviticus, V. N. Livshits. Moscow, 1972. 144 p.

[24] Gomes, P. V. Hybrid genetic algorithm for multi-objective transmission expansion planning / P. V. Gomes, J. T. Saraiva // IEEE International Energy Conference (ENERGYCON). Belgium, 2016. April 4–8.

[25] Zhao, H.-S. Optimal computation of the transmission system expansion planning using the branch and bound method / H.-S. Zhao, L. Chen, T. Wu // Asia-Pacific Power and Energy Engineering Conference. 2009.

[26] Choi, J. A method for transmission system expansion planning considering probabilistic reliability criteria. / J. Choi, T. Tran, A. A. El-Keib, R. Thomas, H. S. Oh, R. Billinton // IEEE Trans. Power Syst. 2005. 20(3). P. 1606–1615.

[27] Soleimani, K. Considering FACTS in optimal transmission expansion planning engineering / K. Soleimani, J. Mazloum // Technol. Appl. Sci. Res. 7(5). 2017. P. 1987–1995.

[28] Jabr, R. A. Optimization of AC transmission system planning / R. A. Jabr // IEEE Trans. Power Syst. 28(3). 2013. P. 2779–2787.

[29] Vasin, A. A. A two-node market in conditions of imperfect competition / A. A. Vasin, E. A. Daylova. MTIP. 6:3 (2014). P. 3–31.

[30] Daylova, E. A. Game-theoretic models of forward and network markets of homogeneous goods: dis. ... candidate of Physical and Mathematical Sciences: 01.01.09; defended: 23.12.2014; approved: 17.10.2014 / Ekaterina Alexandrovna Daylova. Moscow, 2014. 116 p.

[31] Vasin, A. Optimization of transmission capacities for multinodal markets / A. Vasin, M. Dolmatova // Procedia Computer Science. 91, 2016. P. 238–244.

[32] Dolmatova, M. S. Game-theoretic models and optimization problems of energy markets: dissertation ... Candidate of Physical and Mathematical Sciences: 01.01.09; defended: 26.06.2017; approved: 05.04.2017 / Marina Stanislavovna Dolmatova. Moscow, 2017. 95 p.

[33] Khachaturov, R. V. Algorithms for maximizing supermodular functions and their application to optimize the grouping of regions in the region / R. V. Khachaturov // J. calculation. matem. and math. phys. 39:1 (1999). P. 33–44; Comput. Math. Math. Phys. 39:1 (1999). P. 29–39.

[34] Khachaturov, R.V. Basic properties of cube lattices, algorithms for their construction and possibilities of application in discrete optimization / R.V. Khachaturov // J. calc. matt. and mathematical phys. 2015. Vol. 55. No. 1. P. 121–134.

[35] Cherenin, V. P. The solution of some combinatorial optimal planning problems by the method of sequential calculations / V. P. Cherenin // Scientific and methodological materials of the economic and mathematical seminar of the LEMM Academy of Sciences of the USSR. Moscow, 1962. Vol. 2.

[36] Khachaturov, V. R. Mathematical methods of regional programming / V.R.Khachaturov. M., 1989. 297 p.

[37] Carmanov, V. G., Mathematical programming / V. G., Carmanov. Moscow: Nauka, 1986.

[38] Gary, M. Computing machines and intractable problems / M. Gary, D. Johnson. Moscow: Mir, 1982.

[39] Kantorovich, L. V. Application of mathematical methods in the analysis of cargo flows / L. V. Kantorovich, M. K. Gavurin // Compendium of articles Problems of improving the efficiency of transport. USSR Academy of Sciences. 1949. P. 110–138.

[40] Guisewite, G. M. Minimum concave-cost network flow problems: applications, complexity, and algorithms / G. M. Guisewite, P. M. Pardalos // Ann. Oper. Res. 25(1). 1990. P. 75–99.

[41] Vasin, A. A. Optimization of the energy market transport system / A. A. Vasin, O. M. Grigorieva, N. I. Tsyganov // Reports of the Academy of Sciences. 2017. Vol. 475. No. 4. P. 377–381.

[42] Vasin, A. A. Optimization of Transmission Systems for Chain-Type Markets / A. A. Vasin, N. I. Tsyganov // Сборник «Communications in Computer and Information Science, серия Optimization and Applications. OPTIMA 2018», Springer International Publishing Cham, 2019. Vol. 974, P. 482–493.

[43] Vasin, A. A. A model for optimization of transport infrastructure for some homogeneous goods markets / A. A. Vasin, O. M. Grigoryeva, N. I. Tsyganov // Journal of Global Optimization. 2020. Vol. 76. № 3. P. 499–518.

[44] Vasin, A. A. Energy markets: Optimization of transmission networks / A. A. Vasin, O. M. Grigoryeva, N. I. Tsyganov // International Journal of Public Administration. 2019. Vol. 42. № 15. P. 1311–1322.

[45] Tsyganov, N. I. Methods for assessing the prospects of gasification of the country's regions / N. I. Tsyganov, A. A. Vasin // Procedia computer science. 2022. Vol. 214. P. 883–891.

[46] Tsyganov, N. I. Optimization of the transport system for «chain»-type markets / N. I.Tsyganov // Compendium of abstracts «Sociophysics and socioengineering 2018». Proceedings of the Second All-Russian Interdisciplinary Conference, May 23-25, 2018, Moscow: IMP RAS. P. 373–374.

[47] Vasin, A. A. Energy markets: optimization of transportation system / A. A. Vasin, O. M.Grigoryeva, N. I. Tsyganov // Compendium «IX Moscow International Conference on Operations Research (ORM2018)». Moscow: OOO «MAX Press», October 22-27, 2018 Vol. 1, P. 183–190.

[48] Tsyganov, N. I. Optimization of the transport network for the «star-type» energy market / N. I. Tsyganov // Compendium «Tikhonov readings»: scientific conference: abstracts of reports. Moscow: «MaxPress» LLC. 2019. P. 69.

[49] Vasin, A. A. Algorithms for optimizing the transmission network of some energy markets / A. A. Vasin, N. I. Tsyganov // Compendium «Lomonosov readings-2020». Theses. Section «Computational Mathematics and Cybernetics», Series «Section of Computational Mathematics and Cybernetics». Moscow: Publishing House of Moscow University (Moscow), P. 43–45.

[50] Tsyganov, N. I. Algorithms for optimizing the transport structure of the energy market of the «tree»-type / N. I. Tsyganov, I. I. Silaev // Compendium «Tikhonov readings»: scientific conference: abstracts of reports. Moscow: «MaxPress» LLC. 2021. P. 109.

[51] Vasin, A. A. Game theory and models of mathematical economics / A. A. Vasin, V. V. Morozov. Moscow: «MaxPress» LLC. 2005.

[52] Stoft, S. Power System Economics: Designing Markets for Electricity / S. Stoft. N. Y.: Wiley, 2002.

[53] Kleinberg, J. Algorithm Design / J. Kleinberg, E. Tardos. L.: Pearson Education. 2006.

[54] Vasiliev, F. P. Numerical methods for solving extreme problems / F. P. Vasiliev. Moscow: Nauka, 1988.

[55] Sukharev, A. G., Course of optimization methods / A. G., Sukharev, A. V. Timokhov, V. V. Fedorov. Moscow: Nauka, 1985.

[56] Silaev, I. I. Methods of optimization of the transport structure of the energy market of the «tree» type: final qualification work (master's thesis) / Ivan I. Silaev. Lomonosov Moscow State University. Faculty of Computational Mathematics and Cybernetics. Department of Operations Research. Scientific supervisor: A. Vasin. Moscow, 2022. 32 p.

[57] Yushkov, I. R. Development and operation of oil and gas fields: method. textbook / I. R. Yushkov, G. P. Khizhniak, P. Yu. Ilyushin. Perm: Publishing House of Perm. National Research. Polytech. University, 2013. 177 p.

[58] Minkhanov, I. F. Development of oil and gas fields: method. textbook / I. F. Minkhanov, S. A. Dolgikh, M. A. Bartholomew. Kazan: KFU, 2019. 96 p.

[59] Dake, L. P. Fundamentals of the development of oil and gas fields - Fundamentals of Reservoir Engineering / L. P. Dyke; ed. E. M. Simkin; translated from English. B. L. Falaleev. Moscow: «Premium Engineering», 2009. 572 p.

[60] Zakirov, S. N. Design and development of gas fields / S. N. Zakirov, B. B. Lapuk. M.: Nedra. 1974. 376 p.

[61] Margulov, R. D. System analysis in the long-term planning of gas production / R. D.Margulov, V. R. Khachaturov, A. V. Fedoseev. Moscow: Nedra, 1992. 287 p.

[62] Bryukhanov, O. N. Gasified boiler units: Textbook / O. N. Bryukhanov, V. A. Kuznetsov. Moscow: INFRA-M, 2005. 392 p.

[63] Salovaar, J. Coal to natural gas fuel switching and CO2 emissions reduction // Harvard College. 2011. 93 p.

Digital resources

[64] The head of the Ministry of Natural Resources said when Russia will run out of oil and gas // RBC. Business. 11.05.2021. - [Digital resource]. URL: https://www.rbc.ru/business/11/05/2021/609971fe9a7947e065f63cd4 (date of reference 01.05.2022).

[65] Novak estimated gas and oil reserves in Russia // RIA Novosti. 27.02.2019. - [Digital resource]. URL: https://ria.ru/20190227/1551393861.html (date of reference 01.05.2022).

[66] Fuel and Energy Complex of Russia - 2019: stat. compendium// Analytical Center under the Government of the Russian Federation. Issue June. 2020. 62 p. - [Digital resource]. URL: https://ac.gov.ru/uploads/2-Publications/TEK_annual/TEK.2019.pdf (date of reference 01.05.2022).

[67] Renewables 2021 // Global status report. 2021. 371 p. - [Digital resource]. URL: https://www.enerdata.net/about-us/case-study/renewable-global-status-ren21-2021.pdf (date of reference 01.05.2022).

[68] Renewables 2011 // Global status report. 2021. 116 p. - [Digital resource]. URL: https://www.ren21.net/Portals/0/documents/Resources/GSR2011_FINAL.pdf (date of reference 01.05.2022).

[69] World energy outlook 2021 // International energy agency. 2021. 386 p. - [Digital resource]. URL: https://iea.blob.core.windows.net/assets/88dec0c7-3a11-4d3b-99dc-8323ebfb388b/WorldEnergyOutlook2021.pdf (date of reference 01.05.2022).

[70] World nuclear performance report 2021 // World nuclear association. 2021. 68 p. - [Digital resource]. URL: https://www.world-nuclear.org/getmedia/891c0cd8-2beb-4acf-bb4b-552da1696695/world-nuclear-performance-report-2021.pdf.aspx (date of reference 01.05.2022).

[71] Advantages of nuclear energy // Rosatom. About the nuclear industry. - [Digital resource]. URL: https://www.rosatom.ru/about-nuclear-industry/preimushchestva-atomnoy-energetiki/ (date of reference 01.05.2022).

[72] The head of Rosatom spoke about the world's uranium reserves for nuclear power plants // Digital Information Agency «Prime». Energy. 11.11.2021. - [Digital resource]. URL: https://1prime.ru/energy/20211111/835194106.html (date of reference 01.05.2022).

[73] Database of indicators of municipalities. Federal State Statistics Service. - [Digital resource]. URL: https://www.gks.ru/dbscripts/munst/ (date of reference 01.05.2022).

[74] The socio-economic situation of the subjects of the Russian Federation. Federal State Statistics Service. - [Digital resource]. URL: https://rosstat.gov.ru/storage/mediabank/sep_region1.htm (date of reference 01.05.2022).

[75] The balance of energy resources for 2019 - [Digital resource]. URL: https://rosstat.gov.ru/storage/mediabank/2EuHouQF/en_balans2019.htm (date of reference 01.05.2022).

[76] Putin took part in the launch ceremony of the first line of the Amur Gas Processing Plant // RBC. Business. 09.06.2021. - [Digital resource]. URL: https://www.rbc.ru/rbcfreenews/60c0a7c89a7947613db96bd3 (date of reference 01.05.2022).

[77] The cost of construction of the «Power of Siberia» gas pipeline exceeded 1 trillion rubles. // RBC. Business. 27.04.2018. - [Digital resource]. URL: https://www.rbc.ru/business/27/04/2018/5ae344799a794785d0b58ea9? (date of reference 01.05.2022).

[78] The «Power of Siberia» gas pipeline is the largest gas transportation system in Eastern Russia. - [Digital resource]. URL: https://www.gazprom.ru/projects/power-of-siberia/ (date of reference 01.05.2022).

[79] «Gazprom» has proposed to increase the tariff for pumping competitors' gas // RBC. 28.06.2017. - [Digital resource]. URL: https://quote.rbc.ru/news/article/5ae098132ae5961b67a1b2a8? (date of reference 01.05.2022).

[80] The official portal of the Unified Gasification Operator of the Russian Federation. - [Digital resource]. URL: https://connectgas.ru (date of reference 01.05.2022).

[81] The real price for a cubic meter is now known: official data // Southern federal. 01.06.2016. - [Digital resource]. URL: https://u-f.ru/News/economics/u28/2016/06/01/101063? (date of reference 01.05.2022).

[82] The Chayandinskoye field is a resource base for the «Power of Siberia» gas pipeline. - [Digital resource]. URL: https://www.gazprom.ru/projects/chayandinskoye/ (date of reference 01.05.2022).

[83] The estimate of the permanent population as of January 1, 2022 and the average for 2021 (people). Federal State Statistics Service. - [Digital resource]. URL: https://rosstat.gov.ru/storage/mediabank/Popul_Comp2022_site.xls (date of reference 01.05.2022).

[84] The «Power of Siberia»: a strategic project. Infographics. December 2, 2019 - [Digital resource]. URL: https://www.gazprom.ru/press/media/2019/962535/ (date of reference 01.05.2022).

[85] The programme of gasification of the regions of Russia 2021-2025. - [Digital resource]. URL: https://www.gazprommap.ru/program/ (date of reference 01.05.2022).

[86] Igor Kobzev and Alexey Miller signed an updated programme for the development of gas supply in Irkutsk Oblast// IIA «Baikal». 20.06.2022. - [Digital resource]. URL: http://www.38rus.com/more/87066 (date of reference 22.06.2022).

[87] Vostsibugol will supply 19.7 million tons of coal for Irkutskenergo in 1.5 years // Baikal24. 30.11.2010. - [Digital resource]. URL: https://baikal24.ru/text/30-11-2010/197/ (date of reference 01.05.2022).

[88] Carbon Taxes in Europe - [Digital resource]. URL: https://taxfoundation.org/data/all/eu/carbon-taxes-in-europe-2022/ (date of reference 04.11.2023).

List of illustrations

List of tables

Appendixes

Table 11. Characteristics of transport lines of the Irkutsk Oblast natural gas market (part 1)

Line number	Number of the first node	Number of the second node	Shortest distance between nodes, km	l (line length, km)
1	1	61	92	115
2	61	60	53	66
3	60	59	199	249
4	59	58	102	128
5	58	2	125	156
6	2	57	176	220
7	57	26	256	320
8	2	33	223	279
9	58	37	167	209
10	59	35	18	23
11	61	32	71	89
12	1	28	69	86
13	60	46	97	121
14	46	62	105	131
15	62	38	79	99
16	38	12	40	50
17	62	63	69	86
18	63	64	47	59
19	64	16	1.6	2
20	16	13	0.683	0.854
21	64	22	10	13
22	22	45	77	96
23	63	5	198	248
24	5	65	18	23
25	65	17	0.736	0.92
26	65	7	3.7	4.625
27	7	15	9.1	11
28	15	66	25	31
29	66	67	63	79
30	67	49	99	124
31	67	42	149	186
32	66	68	48	60
33	68	27	32	40
34	68	69	106	133
35	69	21	75	94
36	21	70	19	24

Table 12. Characteristics of transport lines of the Irkutsk Oblast natural gas market (part 2)

Line number	**Number of the first node**	**Number of the second node**	**Shortest distance between nodes, km**	l **(line length, km)**
37	70	43	42	53
38	70	39	78	98
39	69	71	19	24
40	71	36	13	16
41	71	72	6.3	7.875
42	72	25	72	90
43	25	47	83	104
44	72	3	54	65
45	3	73	11	14
46	73	18	6.6	8.24
47	18	30	38	48
48	73	74	48	60
49	74	54	26	33
50	74	75	34	43
51	75	51	9.4	12
52	75	76	32	40
53	76	11	2.4	3
54	11	77	1.6	2
55	77	23	2.9	3.625
56	77	48	68	85
57	48	29	74	93
58	76	78	13	16
59	78	20	5.3	6.625
60	78	79	16	20
61	79	53	42	53
62	53	55	51	64
63	53	56	72	90
64	56	52	50	63
65	52	80	65	81
66	80	40	44	55
67	80	34	111	139
68	79	10	20	25
69	10	81	15	19
70	81	44	46	58
71	81	8	21	26
72	8	9	8.4	11
73	9	24	23	29
74	9	82	27	34
75	82	19	0.603	0.754
76	19	31	44	55
77	82	4	7.1	8.875
78	4	6	10	13
79	6	83	22	28
80	83	50	8.5	11
81	83	41	41	51
82	41	14	36	45

Table 13. Results of optimization of the transport system of Irkutsk Oblast natural gas market for various scenarios (production volumes, thousand t.c.f./year)

<table>
<tr><th>Node number</th><th>Producer</th><th>Scenario 1</th><th>Scenario 2</th><th>Scenario 3</th><th>Scenario 4</th><th>Scenario 5</th><th>Scenario 6</th><th>Scenario 7</th></tr>
<tr><td>1</td><td>Kovyktinskoye field</td><td colspan="4">0</td><td>2 732.9</td><td>7 750.4</td><td>8 023.5</td></tr>
<tr><td>2</td><td>Chayandinskoye field</td><td colspan="7">0</td></tr>
</table>

Table 14. Results of optimization of the transport system of Irkutsk Oblast natural gas market for various scenarios (consumption volumes, thousand t.c.f./year)

<table>
<tr><th>Node number</th><th>Consumer</th><th>Scenario 1</th><th>Scenario 2</th><th>Scenario 3</th><th>Scenario 4</th><th>Scenario 5</th><th>Scenario 6</th><th>Scenario 7</th></tr>
<tr><td>3</td><td>Novo-Ziminskaya TPP (NZTPP)</td><td colspan="5">0</td><td colspan="2">564.0</td></tr>
<tr><td>4</td><td>Novo-Irkutskaya TPP (NITPP)</td><td colspan="5">0</td><td colspan="2">1 286.0</td></tr>
<tr><td>5</td><td>TI&TS of the Irkutsk TPP-6 (TPP-7)</td><td colspan="4">0</td><td colspan="3">217.0</td></tr>
<tr><td>6</td><td>Shelekhovsky section of Novo-Irkutskaya TPP (TPP-5)</td><td colspan="5">0</td><td colspan="2">119.8</td></tr>
<tr><td>7</td><td>Irkutsk TPP-6</td><td colspan="4">0</td><td colspan="3">518.0</td></tr>
<tr><td>8</td><td>Irkutsk TPP-9</td><td colspan="5">0</td><td colspan="2">1 404.9</td></tr>
<tr><td>9</td><td>Irkutsk TPP-10</td><td colspan="5">0</td><td colspan="2">1 164.8</td></tr>
<tr><td>10</td><td>Irkutsk TPP-11</td><td colspan="5">0</td><td colspan="2">340.7</td></tr>
<tr><td>11</td><td>Irkutsk TPP-12</td><td colspan="5">0</td><td colspan="2">67.6</td></tr>
<tr><td>12</td><td>Irkutsk TPP-16</td><td colspan="6">0</td><td>97.0</td></tr>
<tr><td>13</td><td>Ust-Ilimskaya TPP</td><td colspan="4">0</td><td colspan="3">438.6</td></tr>
<tr><td>14</td><td>TPP LLC «Teplosnabzhenie» (Baikal TPP)</td><td colspan="7">0</td></tr>
<tr><td>15</td><td>TPP of JSC «Ilim Group» in Bratsk</td><td colspan="4">0</td><td colspan="3">919.7</td></tr>
<tr><td>16</td><td>TPP of JSC «Ilim Group in Ust-Ilimsk</td><td colspan="4">0</td><td colspan="3">639.5</td></tr>
<tr><td>17</td><td>Bratsk</td><td colspan="5">0</td><td>9.1</td><td>32.6</td></tr>
<tr><td>18</td><td>Zima</td><td colspan="6">0</td><td>18.9</td></tr>
<tr><td>19</td><td>Irkutsk</td><td colspan="5">0</td><td>32.7</td><td>157.7</td></tr>
<tr><td>20</td><td>Svirsk</td><td colspan="5">0</td><td colspan="2">27.6</td></tr>
<tr><td>21 – 22</td><td>Tulun, Ilimsk</td><td colspan="7">0</td></tr>
<tr><td>23</td><td>Cheremkhovo</td><td colspan="6">0</td><td>7.4</td></tr>
<tr><td>24 – 37</td><td>Angarsk district, Balagansky district, Bodaibinsky district, Bratsky district, Zhigalovsky district, Zalarinsky district, Ziminsky district, Irkutsk district, Kazachinsko-Lensky district, Katanga district, Kachugsky district, Kirensky district, Kuytunsky district, Mamsko-Chuisky district</td><td colspan="7">0</td></tr>
<tr><td>38</td><td>Nizhneilimsky district</td><td colspan="6">0</td><td>0.535</td></tr>
<tr><td>39 – 45</td><td>Nizhneudinsky district, Olkhonsky district, Slyudyansky district, Taishet district, Tulunsky district, Usolsky district, Ust-Ilimsky district</td><td colspan="7">0</td></tr>
<tr><td>46</td><td>Ust-Kutsky district</td><td colspan="4">0</td><td>0.056</td><td>0.295</td><td>0.905</td></tr>
<tr><td>47 – 56</td><td>Ust-Udinsky district, Cheremkhovsky district, Chunsky district, Shelekhovsky district, Alarsky district, Bayandaevsky district, Bokhansky district, Nukutsky district, Osinsky district, Ehirit-Bulagatsky district</td><td colspan="7">0</td></tr>
<tr><td>Total</td><td></td><td colspan="4">0</td><td>2 732.9</td><td>7 750.4</td><td>8 023.5</td></tr>
</table>

Table 15. Results of optimization of the transport system of Irkutsk Oblast natural gas market for various scenarios (flows, thousand t.c.f./year)

Line	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**	**Scenario 5**	**Scenario 6**	**Scenario 7**
1 (1 → 61), 2 (61 → 60), 13 (60 → 46)				0	2 732.9	7 750.4	8 023.5
3 (60 → 59), 4 (59 → 58), 5 (58 → 2), 6 (2 → 57), 7 (57 → 26), 8 (2 → 33), 9 (58 → 37), 10 (59 → 35), 11 (61 → 32), 12 (1 → 28)							0
14 (46 → 62)				0	2 732.9	7 750.1	8 022.6
15 (62 → 38)						0	97.6
16 (38 → 12)						0	97.0
17 (62 → 63)				0	2 732.9	7 750.1	7 925.0
18 (63 → 64), 19 (64 → 16)				0			1 078.1
20 (16 → 13)				0			438.6
21 (64 → 22), 22 (22 → 45)							0
23 (63 → 5)				0	1 654.7	6 672.0	6 846.9
24 (5 → 65)				0	1 437.7	6 455.0	6 629.9
25 (65 → 17)					0	9.0	32.6
26 (65 → 7)				0	1 437.7	6 445.9	6 597.2
27 (7 → 15)				0	919.7	5 927.9	6 079.2
28 (15 → 66), 32 (66 → 68), 34 (68 → 69), 39 (69 → 71), 41 (71 → 72), 44 (72 → 3)					0	5 008.2	5 159.5
29 (66 → 67), 30 (67 → 49), 31 (67 → 42), 33 (68 → 27), 35 (69 → 21), 36 (21 → 70), 37 (70 → 43), 38 (70 → 39), 40 (71 → 36), 42 (72 → 25), 43 (25 → 47)							0
45 (3 → 73)					0	4 444.2	4 595.5
46 (73 → 18)						0	18.9
47 (18 → 30)							0
48 (73 → 74), 50 (74 → 75), 52 (75 → 76)					0	4 444.2	4 576.6
49 (74 → 54), 51 (75 → 51)							0
53 (76 → 11)					0	67.6	75.0
54 (11 → 77), 55 (77 → 23)						0	7.4
56 (77 → 48), 57 (48 → 29)							0
58 (76 → 78)					0	4 376.6	4 501.7
59 (78 → 20)					0	27.6	27.6
60 (78 → 79), 68 (79 → 10)					0	4 348.9	4 474.0
61 (79 → 53), 62 (53 → 55), 63 (53 → 56), 64 (56 → 52), 65 (52 → 80), 66 (80 → 40), 67 (80 → 34)							0
69 (10 → 81), 71 (81 → 8)					0	4 008.2	4 133.3
70 (81 → 44)							0
72 (8 → 9)					0	2 603.3	2 728.4
73 (9 → 24)							0
74 (9 → 82)					0	1 438.6	1 563.6
75 (82 → 19)					0	32.7	157.7
76 (19 → 31)							0
77 (82 → 4)					0		1 405.9
78 (4 → 6)					0		119.8
79 (6 → 83), 80 (83 → 50), 81 (83 → 41), 82 (41 → 14)							0