JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2010

(*Contents Continued on Back Cover*)

ISSN 1509-4553 on-line: ISSN 1899-8852

Warsaw 2010

Circulation: 300 copies

Sowa - Druk na życzenie, www.sowadruk.pl, tel. 022 431-81-40

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

This issue covers diverse themes related either to socio-economic impacts of telecommunication and Internet, to techno-economic models of telecommunication development and markets, or to game-theoretical modeling of related subjects. It contains ten papers.

Andrzej P. Wierzbicki in his paper *Socio-Economic Challenges from and for Future Internet* takes the view that from the dynamics of the development of Internet itself (or Future Internet) result challenges for the socio-economic development and even for the paradigm of understanding economy; such challenges must be taken into account first, then – in a feedback loop – they lead to the modification of socio-economic challenges for Future Internet. This change of perspective is necessary because the analysis of socio-economic challenges for Future Internet is typically based on a classical paradigm of equilibrium economics. While useful, this paradigm is limited and has been shown incomplete, e.g., by the recent worldwide financial and economic crisis. This paper shows that the informational revolution – including but not limited to the development of Internet – has already changed the behavior of main socio-economic agents as well as of economic markets and our understanding of them; thus, new perspectives are needed. The paper also shows that a useful perspective is that not of expected benefits from Future Internet – there are many such benefits and they will surely motivate enough economic development of Future Internet – but of socio-economic threats resulting from the dynamic development of the Internet. There are several such threats and conflicts discussed in the paper.

Cristina Casanueva-Reguart and Antonio Pita S. in the paper *Telecommunications, Universal Service and Poverty in Mexico: a Public Policy Assessment (1990–2008)* analyze the design and implementation of telecommunications service policies targeted at the poorest regions of Mexico (1990–2008). It begins by defining universal access and service policies, their economic and social rationale. Secondly, it discusses the scope of public policies on universal service provision designed by Mexican authorities to achieve the goal of universal access. Thirdly, the paper analyzes the distributive effects of this set of policies among the poorest sectors of the population. The sources on which this research was based were two national surveys: the *Household Income and Expenditure Survey* (2008), and the *Household Survey of the Access and Use of Information Technologies* (2007). The additional information on regional economic development was based on the poverty indexes by the national population council and economic information given by Mexico Census Bureau. Additional use was made of the annual reports prepared by Ministry of Communications, statistics published by the Federal Telecommunications Commission and official documents prepared by the government agencies. Finally, a series of in-depth interviews was conducted with the former representatives of the Office of Rural Telephony. Finally, the article discusses, in the light of available evidence, possible explanations for the apparent failure of the universal service policy that was implemented to bring at least basic voice services to Mexico's neediest. This explanation is seen in profit-maximizing policies of privatized telecom operators, not interested in providing universal service; to overcome this, a stronger regulation policies are needed. of available evidence, possible explanations for the apparent failure of the universal service policy that was implemented to bring at least basic voice services to Mexico's neediest. This explanation is seen in profit-maximizing policies of privatized telecom operators, not interested in providing universal service; to overcome this, a stronger regulation policies are needed.

Yousef Gasiea, Margaret Emsley, and Ludmil Mikhailov in the paper *Rural Telecommunications Infrastructure Selection Using the Analytic Network Process* analyze the applicability of a multicriteria decision-making method, specifically the analytic network process (ANP), to model the selection of an appropriate telecommunications infrastructure technology, capable of deploying e-services in rural areas of developing countries. It aims to raise awareness among telecommunication planners about the availability of ANP, and to demonstrate its suitability to enhance the selection process. The proposed model is constructed based on concerned experts' subjective views of relevant selection criteria and potential technology alternatives. Its network structure takes account of all possible dependencies and interactions among criteria and alternatives.

Paweł Olender in his paper *Stochastic Models in Techno-Economic Analysis of Broadband Access Networks* presents a techno-economic model of broadband access, which was implemented and used to carry out analyses for one of the biggest city in Poland. A stochastic approach was applied in order to take uncertainty into consideration, resulting in a more robust solution, therefore, improving the safety of investment. Analyses concern FTTH (fibre to the home) technology, a type of generic FTTx network architecture. It uses optical fibre in local telecommunication loop, what is becoming more and more popular. Presented results show the usefulness of techno-economic surveys in planning access networks development. The appropriate choice of network parameters, such as the aggregation ratio, is essential and could significantly influence the profitability of investment.

Wojciech Stańczuk, Piotr Pałka, Jozef Lubacz, and Eugeniusz Toczyłowski in their paper *A Framework for Evaluation of Communication Bandwidth Market Models* present a method of analysis of market-based models for resource allocation in communication networks. It consists of several stages: classification of a market model, generation of input data, data adaptation to a tested model, test calculations and, finally, presentation and interpretation of results. A set of general criteria to assess various models has been proposed. Tests are run using dedicated computer applications, data is stored in open XML-based format originated in the multicommodity market model. Network topologies are derived from the SNDlib library.

Piotr Pałka and Eugeniusz Toczyłowski in their paper *Pricing Rules Comparison in the Context of Bandwidth Trade* compare two pricing rules in the context of bandwidth trade. Allocation and pricing rules, together with a set of signals received from independent agents, constitute a market mechanism. The analyzed pricing rules are the well known Vickrey-Clarke-Groves rule (VCG) and the parametric pricing rule (PPR). These pricing rules are applied to the allocation rule specified by the balancing communication bandwidth trade model (BCBT).

Kamil Kołtyś, Piotr Pałka, Eugeniusz Toczyłowski, and Izabela Żołtowska in their paper *Bandwidth Trading: A Comparison of the Combinatorial and Multicommodity Approach* compare two different double-sided bandwidth auction mechanisms, that seem to be well suited approaches for trading indivisible units of bandwidth: combinatorial auction c-SeBiDA and multicommodity mechanism BACBR-I. The c-SeBiDA mechanism considers two types of commodities: inter-node links and paths consisting of particular links. Market participants may bid a single link, or a bundle of links, constituting a specific path. The BACBR-I mechanism is a multicommodity exchange model, that allows bidders to place buy offers not only for individual or bundled links, but rather for end-to-end connections. Therefore, it is the decision model that allocates the most efficient links to connections. A large set of experiments was performed to test the allocation and computational efficiency obtained under both approaches.

Krzysztof Malinowski, Ewa Niewiadomska-Szynkiewicz, and Przemysław Jaskóła in their paper *Price Method and Network Congestion Control* consider price instruments that are useful in achieving market balance conditions in various markets. Those instruments can be also used for control of other composite systems. The formulation and basic properties of the price method are reviewed and then the congestion control by price instruments in a computer network is described and tested.

Keiichi Niwa, Tomohiro Hayashida, and Masatoshi Sakawa in their paper *Computational Methods for Two-Level 0-1 Programming Problems through Distributed Genetic Algorithms* consider a two-level 0-1 programming problem in which there is not coordination between the decision maker (DM) at the upper level and the decision maker at the lower level. A revised computational method is proposed that solves problems related to computational methods for obtaining the Stackelberg solution. Specifically, in order to improve the computational accuracy of approximate Stakelberg solutions and shorten the computational time of a computational method implementing a genetic algorithm (GA) proposed by the authors, a distributed genetic algorithm is introduced with respect to the upper level GA, which handles decision variables for the upper level DM. Parallelization of the lower level GA is also performed along with parallelization of the upper level GA. The proposed algorithm is also improved in order to eliminate unnecessary computation during operation of the lower level GA, which handles decision variables for the lower level DM. In order to verify the effectiveness of the proposed method, comparisons with existing methods are outlined by performing numerical experiments to verify both the accuracy of the solution and the time required for the computation.

Ichiro Nishizaki, Tomohiro Hayashida, and Noriyuki Hara in the paper *Coordination Games with Communication Costs in Network Environments* deal with a coordination game in a network where a player can choose both an action of the game and partners for playing the game. In particular, a player interacts with players connecting through a path consisting of multiple links as well as with players directly connecting by a single link. Decay or friction of payoffs with distance as communication costs are represented, and the effect of the communication cost on the behavior of players in the game and network formation is analyzed. Properties of equilibrium networks are investigated by classifying the link cost and the communication cost, and the diversity of the equilibrium networks is indicated.

Andrzej P. Wierzbicki
Guest Editor

Paper

Socio-Economic Challenges from and for Future Internet

Andrzej P. Wierzbicki

Abstract—There are several papers available addressing the challenges *for* Future Internet that result from socio-economic aspects; such challenges must be obviously taken into account when constructing and developing Future Internet. This paper, however, takes an opposite view that *from* the dynamics of the development of Internet itself (or Future Internet) result challenges for the socio-economic development and even for the paradigm of understanding economy; such challenges must be taken into account first, then – in a feedback loop – lead to the modification of socio-economic challenges *for* Future Internet. This change of perspective is necessary because the analysis of socio-economic challenges for Future Internet is typically based on a classical paradigm of equilibrium economics. While useful, this paradigm is limited and has been shown incomplete, e.g., by the recent world-wide financial and economic crisis. This paper shows that the informational revolution – including but not limited to the development of Internet – has already changed the behavior of main socio-economic agents as well as of economic markets and our understanding of them; thus, new perspectives are needed. The paper also shows that a useful perspective is that not of expected *benefits* from Future Internet – there are many such benefits and they will surely motivate enough economic development of Future Internet – but of socio-economic *threats* resulting from the dynamic development of the Internet. There are several such threats and conflicts that can be foreseen: the conflict between corporatization and governance; the conflict between direct and indirect limits to freedom; the trend towards elitarism inherent in Internet development dynamics versus democracy; the threat of network and computer domination over people, etc. Some of such threats and conflicts are discussed in the paper.

Keywords—Future Internet, informational revolution, socio-economic changes and challenges.

1. Introduction

The beginnings of informational revolution can be dated not from the development of computers, but from developments that enabled broad social use of computers and computer networks, thus from the development of an inexpensive personal computer (Apple II, 1977) and from the de-classification of Internet (1983), thus around from 1980. Computers were invented 40–50 years earlier: analog computers in 1931 (by a telecommunication engineer, later presidential advisor Vannevar Bush), digital computers in 1936 (by a telecommunication engineer Konrad Zuse, not by a mathematician Alan Turing, whose theoretical paper was published after the engineering patent of Zuse). This is similar to the dating of industrial revolution in around 1760 with the inventions of James Watt – who only made possible a broad social use of a steam engine (dangerous because unstable before Watt) invented by Newcoman at least 40 years earlier, by supplementing the engine with a feedback control system of rotational speed, thus making it safe for a broad use.

Here I should make clear my basic assumption about the development of new technologies. The popular theories of a co-evolution of social attitudes and technological solutions (see [1], [2]) are applicable only to a continuous near-equilibrium evolution, certainly not applicable to such events as the invention of James Watt, the emergence of Internet, or even the construction of a Future Internet. By this I do not mean that Future Internet should be a *clean-slate* solution; I mean only that a technological solution of this magnitude of social impact was until now – and should be in future – based on a vision that hopefully would meet social expectations at least 40 years from its conception. Internet was based on such a vision, even if its success has outgrown the scope of that vision. Every radically new technology (computers, transistors, mobile telephony, digital television, see [3]) is usually conceived at least 40 years before its broad social use and is based on a vision, not on a co-evolution of social attitudes and technological solutions.

Today it is clear that a broad social use of personal computing and of computer networks (Internet, WWW) has changed essentially the social fabric of developed societies, and has created many new opportunities and challenges (see, e.g., [4], [5]). The informational revolution is manifested in its three main megatrends [6]:

- the technological megatrend of digital integration (also called convergence);
- the social megatrend of dematerialization of work and changing professions;
- the intellectual megatrend of changing perception of the world.

We shall not discuss here these megatrends in detail, I quote them only to illustrate that we can take for granted tremendous socio-economic and even intellectual changes resulting from informational revolution, even if this revolution was enabled by technological developments.

Thus, when the authors of an extremely interesting book *Towards the Future Internet: A European Research Perspective* [7] include several papers addressing the challenges *for* Future Internet that result from today's socio-economic needs, I respond with the question what are the challenges *from* Future Internet, the aspects that the Future Internet will add to the informational revolution and socio-economic developments in the future. To analyze this issue, I will follow the *dynamic programming* paradigm: imagine what would be the world in the year 2050 and then analyze challenges resulting from the dynamics and conflicts of the development. The choice of the date 2050 is substantiated because we see – already from the examples of delays quoted above, but the mobile telephony had a delay of about 50 years before its broad social use, the digital television even more, the transistors about 40 years – that new technologies achieve its full social impacts with substantial delay, amounting today to 40–60 years; this delay might shorten in the future, but not immediately. Therefore, in 2050 we shall count with a broad social use of technologies that we start to develop today. The issue is what social needs they will satisfy and what socio-economic conflicts or threats they will create; to analyze this issue, we need a vision of the year 2050.

2. A Personal Vision of the Year 2050

This is obviously a personal vision, expressing my personal views and experience in future studies. It is true that during the informational revolution everything flows, the world is a collection of chaotic systems from which new patterns of order emerge. Thus, many unpredictable changes can occur and every detailed, quantitative forecast should be treated with suspicion. However, the dynamic of social changes is slow, people do not change their accustomed modes of behavior, the *qwerty* keyboard will be used in 2050 even if it is not optimal.

Moreover, the stories about full unpredictability of the world, about the phenomenon of *black swan*[1], are means of brainwashing people. Internet was not a *black swan*, already in 1970 Arpanet (although classified) started its functioning, protocol IP and e-mail (together with using the sign @ in addressing) were devised in 1972, Internet was de-classified in 1983, and the fact, that after 40 years only about 1/6 of world population uses Internet is a small delay compared to mobile telephony or digital television. Thus, until 2050 we shall certainly have many novel inventions, theories or even scientific revolutions, but they will not have a broad social impact before 2050. Imagine, for example, that somebody invents today an *avio-car* (a flying car), sufficiently efficient and with low emissions. Before it will be developed to a sufficiently inexpensive and safe version (together with appropriate traffic regulations) for a broad social use, certainly more than 40 years will be needed. On the other hand, some inventions or developments known today might be developed for a broad social use before 2050.

[1]*Black swan* is a metaphor of an unpredictable phenomenon (see [8]). The main example of this phenomenon was supposedly the emergence of Internet. Already in early 80-ties, I tried to convince my Polish colleagues about the inevitability of the development of social importance of Internet, only very few believed me. Thus, we have not the phenomenon of *black swan*, rather the phenomenon of *Cassandra*.

Thus, if we guess correctly which rudimental developments or inventions known today will meet in future broad socio-economic needs, we can if not forecast, then at least construct a probable vision of the world in 2050. I use the words *constructing a vision of future*, because humanity always constructed future based on some visions; if we build a house, we construct future following some vision.

In such a vision, I see three main development forces that correspond to main socio-economic needs and will shape the future society. These are:

- *the need of living in a clean environment*, expressed by the idea of *sustainable development*;
- *the need of boundless communication*, expressed by the *informational revolution* with all its derivative consequences;
- *the need of prolonging life*, expressed by the idea of *bio-technological revolution*.

The last one – the bio-technological revolution – is also related to the concept of *radical evolution*, or human evolution reinforced by technology (see, e.g., [9]). However, I do not believe that major social needs will contribute to the start of radical evolution, or full bio-technological revolution (with similar or even larger controversies than those associated with the information revolution today), before 2100. Until 2050, on the other hand, the need of prolonging life will support a broad implementation of some elements of bio-technological revolution, particularly for elder people.

The second one – informational revolution – started around 1980, as we noted above, and already has tremendous impacts. It will continue; together with other main development forces, it will determine the socio-economic impact *from* the Future Internet.

The first one – sustainable development – expresses a major social need of living clean and preserving environment for future generations. It was perceived earlier than the other major needs (see, e.g., [10]), thus the concept of sustainable development is well known and broadly discussed. However, I believe that the problem how and in what proportions we should support development caring about its sustainability and clean environment at least in the interest of our children will remain a fundamental one at least until 2050 and will determine the solutions of related problems such as energy provision, transportation, life style, details of environmental protection. Both the informational revolution and the rudiments of bio-technological revolution will contribute to the solution of such problems.

Before turning to the issue of challenges *from* the Future Internet, we shall analyze shortly what might be the impact of these three major socio-economic forces before 2050. However, I must stress – before this analysis – a fundamental assumption: until 2050 we will not be faced by a major global catastrophe, economic larger than the current crisis, military of world-wide character, or cosmic. This assumption is obviously optimistic, and it can be fulfilled only if we increase *global governance*; thus, I will also discuss shortly this issue.

2.1. Sustainable Development

Without a major catastrophe, the world economy will be forced – by public opinion and preferences of consumers – to take into account increasingly more demands of sustainable development, to limit harmful emissions, etc. This is fully possible when using technologies known rudimental today – it requires "only" money and time for their detailed development, which will not happen spontaneously, because free market promotes technologies that bring fast returns, is bad in long term rationality. Therefore, we will face slow but inevitable restriction of free market by the character of demand, but also norms and regulations imposed by "green" consumers.

Even today, there is an intensive research on diverse technologies either of car engines that are environmental friendly, or of limiting harmful emissions from power plants, foundries and chemical factories, or of alternative energy sources. In 40 years, many new inventions accelerating this change will be made, but decisive will be the confrontation of short-term interests of corporations and long-term interests of humanity, leading to slow and gradual, but inevitable sharpening of environmental norms and regulations. It might appear that the interests of consumers and entire humanity are less strongly represented, are doomed to loose in opposition to the strength of large corporations; but the history of last 40 years shows the opposite. If children in schools are taught to consider environmental protection as a higher value, then as young consumers they will not buy products of corporations that do not show sufficient environmental care, and a corporation might end as, e.g., in the case of Chrysler.

It might be optimistic, but I believe that diverse sources of environmental pollution will be until 2050 several times reduced – at least, in developed countries, but the developing countries will catch up with environmental protection, even if with more difficulty. I am not sure that this will be sufficient to preserve the natural environment on our planet in an acceptable state, but it gives a chance. Nevertheless, it will be a difficult process, with many controversies and consequences. This process requires using new technologies, increasing automation and robotization, supports and is supported by the transition towards knowledge economy, but on the other hand it means also the dematerialization of work, thus global escape of some industries towards developing countries, social disorders related to large unemployment during the period of strong structural change – and all next 40 years will be such a period. Thus the process of adapting to sustainable development and to new, "green" technologies will not be easy, it will require a permanent re-education of societies.

This is the basic challenge before the society of entire globe: *if sustainable development can be realized only by developing knowledge-based economy, then global education level must continuously increase*. We might expect that in 40 years a condition of employment in developed countries – with the exception of clearly subsidiary, supporting service work – will be higher education which will be practically universal in developed countries. This does not mean that the proportions of educational profiles will precisely meet the demands of labor markets – just the opposite, we must become accustomed to the idea that a taxi driver with education on the master of management level is not a singularity, only a perturbation of fate and not a social waste – because a highly developed society should be able to support an excess of learning.

The development in this direction will be not uniform around the globe. Countries such as Finland, that devote a large portion of national income to education and science, will win the competition towards sustainable development and knowledge-based economy. Countries such as Poland, where the government systematically cuts the funding of education and science, relying rather on the private efforts of citizens educating themselves mostly in private universities, will be doomed to marginalization and the role of civilization peripheries. They will be overtaken, e.g., by the countries of the Far East, such as China, Korea, Vietnam, who devote much more attention to science and education.

Sustainable development requires also a substantial change of professional proportions in society. This does not mean that in 2050 we shall observe globally the same proportions that we observe today in the USA or Japan. For example, the issue of a large part of society living in villages and from agriculture can be resolved in diverse ways, not only through the reduction of the number of farmers, increase of farm area and the escape of remaining people to slums around big cities. It might be resolved by a redistribution of work and living to rural areas, aided by network technologies and new Internet.

Nevertheless, it is expected that less people will work at the production of food and, at the same time, we can hope that regions of endemic hunger will be eliminated from the globe. Similarly, less people will be employed in industrial production. The rest will be employed in diverse services, such as education, health service and old people care.

Generally, the idea of sustainable development is based on respect for nature. We can have the optimistic hope that until 2050 the global environmental situation will be improved, despite a further increase of the world population (which, according to prognoses of United Nation Organization (UNO) based on the research of International Institute for Applied Systems Analysis (IIASA), in these decades will be slower and will attain a maximum precisely around 2050, with an irreversible increase of the proportion of old people). It will not be easy, an increased international cooperation towards this objective is necessary, helping the developing countries to avoid environmental pollution excesses known to developed countries. But the essential condition is the stress on ecological responsibility in education, creating the domination of green-friendly consumers.

It is also important to understand well the relation between nature and technology: technology in itself does not kill nature[2]. All conflicts concerning environmental protection are in fact not conflicts about technology proper, only about its use – often occurring between short term interests of market entrepreneurs (obviously using technology with harm for nature, if this brings profits) and long term interests of protecting nature, e.g., by local communities.

On the other hand, we observe a slow progress in propagating and understanding environmental values: ecological responsibility is taught world-wide and some large corporations – e.g., in Japan – gradually stress environmental values. We can thus expect that the impact of green-friendly consumers and local communities will gradually extend to global scale and will help in improving global ecology.

However, this will not occur without strengthening global governance, which I discuss separately. Global governance might be also needed for other aspects of protecting global environment. In long history of our planet there were several cosmic catastrophes which changed global environment radically. Humanity might decide that we are rich and wise enough to prevent in future such catastrophes; initial research in this respect is already conducted, but and intensification of such activities would require, e.g., establishing an international base on the Moon with the purpose of observation of approaching cosmic bodies and suitable reactions (changing their trajectories).

2.2. Informational Revolution and Knowledge Society

All structural changes today, closely related to *informational revolution*, transform the economy towards *knowledge-based economy* (see, e.g., [13]), or even *knowledge civilization* (see [14], [15]). Without discussing these issues in detail[3] I must stress that the force of informational revolution will not diminish until 2050 though it might address different aspects. Two examples might be relevant.

Multimedia record and transmission. Social demand for multimedia record of information and occurrences as well as transmission of such records will grow because of diverse reasons, such as increasing interest in enriched films becoming a substitute of books, the necessity of preparing and transmitting multimedia teaching materials in spreading distance education, an increasing demand for multimedia telephony (such as Skype), etc. We must be aware that even if the methods of recording and transmitting distance education materials are highly developed, the tools for creating such materials are not sufficiently developed, standardized and ready for market penetration. Moreover, social customs in this respect might change slowly (e.g., because of attachment of part of society, myself included, to the traditional form of books). However, a change in this direction is inevitable, because of many reasons, such as the power of *Open Access* initiatives that provide networked free access to educational resources including increasingly multimedia forms. On the other hand, we cannot expect universal multimedia character of record and transmission of information until around 2050, because of large delays of social demands in such cases.

Ambient intelligence or wireless sensor networks. These diverse slogans characterize different approaches (in European Union and United States) to the same problem: how to use universally inexpensive computer tools, such as microprocessors equipped with sensors and radio, to provide for intelligent environment in human habitat. The slogan of *ambient intelligence* was put forward around the year 2000 by the Information Society Technology Advisory Group (ISTAG) of the European Commission as a driving engine of European economy. As a member of this group I raised then the objection that the delays and generally slow dynamics of changing social customs make such a slogan unrealistic before 2030 or even 2050. The social resistance in this respect might be large, because even if the needs of health care of older people might demand continuous and non-intrusive monitoring of the identity, presence, consciousness, breath, heart beat, etc., of people in a given room, not everybody would agree to enter such a room without warning about monitoring and recording his personal parameters. As a result, we can expect until 2050 substantial development of ambient intelligence, but not its universal applications. This is related to the wide-spread fear of Orwellian utopia, of using intelligent environment in human habitat for an excessive social control by too am-

[2]This is very badly understood by most humanists, social scientists and even natural scientists, since they usually do not have courses of technology in their curriculae – while technologists attend courses of all these sciences, e.g., of philosophy. As a result, representatives of these sciences perceive technology through the lens of humanistic philosophy of technology which is unable to understand technology at all because of the lack of a direct contact with technology. This often leads to the basic error, justly condemned in cultural anthropology: the error of cultural imperialism, judging a different culture without fully understanding it. For example, some humanist philosophers of technology condemn technology and technocracy without understanding that they actually speak about an aggregated notion including socio-economic use of technology in market economy, not about technology proper. Technology proper is the art of creating tools and artifacts characteristic for a given civilization era (see [11], [12]), and can be used both for good and bad purposes. This great misunderstanding of technology is characteristic for entire 20th Century, starting, e.g., with Albert Einstein who wrote already in 1917 that "The advance of technology is a hatchet in hands of a degenerated criminal" (see, e.g., [9]). Martin Heidegger described the same issue much more deeply saying that the danger is not in the advance of technology itself, but a dangerous fascination with the possibilities of technology by people, particularly by people in power: "man exalts himself [with the possibilities of technology] and appears to be the lord of the world" [11]. The fact that humans would cease to be human if they stopped technology creation escapes the attention of humanists because they have inadequate education in technology.

[3]Knowledge was obviously used as an economic resource in all civilization eras, but now it becomes – first time in the history – a decisive productive resource, dominating labor and capital, as a result of informational revolution. The era of knowledge civilization will be probably not the last in human history, but it will continue at least for the entire 21th Century.

bitious politicians, or even fear of the domination of networks, computers and robots over humanity, which subject I comment in more detail later.

Beside these two examples, there are many other areas of the impact of informational revolution on economy and society. The great *megatrend of digital integration* (or *convergence*) was not yet exploited fully. Between other areas of digital integration, several decades yet will be required, e.g., until diverse media such as newspapers, radio, television, Internet will become more deeply integrated. Around 2050 we might, however, expect a more deep integration of diverse information media, their networked access in a selective or fully integrated form. Since economic and political power of controlling media is tremendous, only this reason – between several others – is sufficient for the development of several new generations of Internet until 2050; today we work intensively on the techniques of *Future Internet*, a main subject of this paper, but we should be aware that there might be several Future Internet versions.

There will be also an inevitable impact of informational revolution on the paradigms of economic science. There are many aspects of this impact, but most important appear to me the *oligopolization of economy* and the *conflict about property of knowledge*. Oligopolization of economy results from the fact that the increasing role of knowledge and intellectual property in production costs inevitably leads to an increase of positive effects of scale and decrease of relative marginal costs; the cost of duplicating a DVD plate is much smaller than the cost of a film production. Therefore, the relation of market price to marginal cost, a paradigmatic foundation of free market theory, is lost in informational revolution era. A possible explanation is the domination of oligopolistic economy (or monopolistic, but this form is tightly regulated). This issue might be studied using classical mathematical game theory to compute how many times an oligopolistic market price without collusions can exceed ideal free market price, given market share and elasticity of demand (and how much actual prices indicate tacit collusions on the oligopolistic market), but for some reasons such investigation is treated as a tabu in neoclassical market economics and oligopoly theory is not taught in detail in economic departments.

The conflict about property of knowledge occurs between a classical equilibrium relation of individual knowledge and the intellectual heritage of humanity on one side, and the new, perturbing the classical equilibrium trend of *corporate privatization of knowledge* – including both the individual knowledge of employees of the corporation and as much of intellectual heritage of humanity as can be privatized by a corporation. This conflict is very serious even today, large corporations do everything to maximally privatize common knowledge of humanity and engage in this respect the neoliberal interpretations of intellectual property rights (see, e.g., [16]). This conflict will probably intensify and might become the basic conflict of knowledge civilization. This conflict is also dangerous, because – in opposition to all classical paradigms of economics – common knowledge is not a degradable good (it usually grows, is not diminished by a common use). Hence the classical argument of *the tragedy of commons* (used to substantiate privatization of common resources)[4] is not applicable to knowledge: *it is better for a society, if as much knowledge as possible remains public property*. This means, however, that knowledge based economy requires a fundamental change of the paradigms of economics.

Even larger, than on economy, is the impact of informational revolution on society, even if a part of this impact is related to economy. The *great megatrend of dematerialization of work* during informational revolution, substituting people by automata and robots in hard productive work, leads to an increase of the share of services in economy and has many positive consequences. Among most important among such consequences might be the creation of material conditions for equality of women (it is the computer and the robot that enable such equality); actual equality of woman might be distant yet because cultural relations and customs change slowly, but around 2050 we might expect that women will achieve globally significant progress in actual realization of their equal rights. However, this megatrend has also obvious negative consequences that must be continuously counteracted: the dematerialization of work leads to a change of professions, disappearance of old ones and emergence of new ones that will continue until and beyond 2050, and results in so called (misnamed)[5] *structural unemployment*. Such unemployment is not a temporary phenomenon, it can be counteracted only by intensive reeducation of labor force, in which distance and electronic education might be decisively used.

Possibly the most important aspect of the social impact of informational revolution is *the annihilation of spatial constraints in the access to information and knowledge and communication between people*, i.e., the gradual spread of multimedia access and network communication. This aspect is possibly more important than the Gutenberg revolution that made the access to information and knowledge universal through books – since books do not fully annihilate the spatial constraint: one has to buy a book and bring it to a small village, or to travel to a great library in a city, while in Internet it is sufficient to have a broadband access from anyplace. Depending on the conclusions of the conflict about the property of knowledge and the success of initiatives such as *Open Access*, after the spatial annihilation might follow at least a partial annihilation of the economic constraints (costs). Multimedia character of the access to information and knowledge, and also multimedia character of communication, might have a positive impact on creativity in using these sources and on the spread of distance and electronic education (see, e.g., [15], [17]). In total, this is a great social revolution which will change not

[4]A common pasture in a local community is degraded by its too intensive use, hence it is better to privatize it.

[5]Structural unemployment implies that the structure of economy has changed and labor force must adapt to this change; but what if the structure is continuously changing due to informational revolution and the speed of change is limited precisely by the speed of adaptation of labor force?

only the conditions of social life and customs, but also will influence the trends of spatial agglomeration and regional policy.

Until 2050 it might come to a reversal of the trend of urbanization of the world, to the beginnings of an actual realization of the idea of *global village*. This idea was originated much too early and until now found only derision between regional economists: all around the world, the trend of urbanization continues. However, people are already tired of everyday long travel between the place of living and the place of work; the larger a metropolitan agglomeration, the more probable are many hours of travel to work. The idea of distance work was also premature, such changes require long time of social adaptation, but until 2050 it might become a social reality. This means that regional policy should seriously consider the possibilities of regional socio-economic development based on the spread of Internet connections, used for the activation of villages and an attraction of out-migration from large cities.

Perhaps the most important social consequences of informational revolution concern already mentioned, fundamental change of educational systems. It is necessary to make universal not only high education, but also continuing education, to large extent realized via distance or electronic education. This must be based on a networked, multimedia access to sources of information and knowledge, must use this type of access to stimulate creativity, prepare new generations for life in a new society. This also means a necessity of changing educational paradigms and of a deep reform of all levels of education, starting with elementary schools. This will be not an easy change, the most paradigmatic or even dogmatic is the educational science that successfully resists all changes. This will not be an inexpensive change, it will probably begin with most developed or most forward-oriented countries (such as, e.g., Finland). Countries that will try to resist or economize on this change will find themselves marginalized.

Another basic aspect of informational revolution is that it enables an improvement of the relation between people and nature. For example, diverse distributed sensors connected in wireless networks can much better monitor the quality of natural environment. Moreover, in knowledge based economy it is easy to promote environment-friendly innovations. This positive feedback between informational revolution and knowledge-based economy on one side and the protection of natural environment on the other side is the reason of my optimism concerning the idea of sustainable development discussed above.

This does not mean that the informational revolution does not bring environmental threats. The main such threat is an excessive exploitation, in a sense over-saturation with signals, of the natural electromagnetic environment of Earth. This does not result as yet in serious dangers for human health, the so called electromagnetic compatibility of electronic equipment is a subject of severe tests and norms, and the electromagnetic spectrum management (allocation of frequencies for commercial and other uses) is an important subject of governmental control. However, diverse other possibilities of utilizing electromagnetic spectrum, such as radio-astronomy, are seriously constrained by the commercial saturation of this spectrum; this is one of the reasons for the necessity of constructing a permanent base on the Moon. Until 2050, the issues of management of electromagnetic spectrum might become an area of socio-economic conflicts, similar to today's environmental conflicts.

However, the main effect of informational revolution in the relation of people and nature is the annihilation of spatial constraints in the access to knowledge and communication discussed above – which might result in a choice of living place in a close contact with nature, not only in a village but also possibly in a forest. Thus, the idea of a *global village* might become actually the *global forest*. This does not mean that until 2050 cities will vanish – just the opposite, they will grow at least until that date. However, until 2050 we might observe the beginnings of the opposite trend, the trend of *global forest*.

2.3. Biotechnological Revolution

Elementary biotechnologies, such as genetically modified crops, have already strong impact on the global economy; it might be also argued that biotechnology is as old as agriculture.[6] However, we are far away from an actual biotechnological revolution, including radical technological changes in human evolution. The speculations about *radical evolution* – the vision of a cyborg as a result of a new, mostly artificial product of biotechnological revolution – are already frequent, but far from realism and such revolution will not occur for sure before 2050, probably also not before 2100. This is because even today we observe a significant social resistance against excessive or unjustified automation of actions customarily reserved for people, against domination of human subjectivity by computers, networks and robots, or generally domination of computer over people. These attitudes will not favor radical biotechnological evolution that inevitably will include implantation of microprocessors into human body; we can expect serious social resistance that will significantly delay radical evolution.

The beginnings of biotechnological revolution and radical human evolution we might observe in areas, where the social demand will be nevertheless substantive: such an area is *health care for elder people*. The implantation of a microprocessor only to stimulate heartbeat or the use of artificially developed bone cells in order to rejuvenate old bones encounters much less social resistance, if it is evidently needed and helps. Such technologies will encounter strong economic demand, which is necessary for their gradual improvement, decrease of costs, universal accessibility. Together with elements of ambient intelligence for the non-intrusive monitoring of the health of old people mentioned

[6] What is the production of beer or whisky, if not a biotechnology?

above, or even with mobile robots taking care of elder people, socially acceptable elements of biotechnological revolution will become a natural enhancement and continuation of informational revolution. However, their broader spread will be restricted to the cases of obvious need and helpfulness. The health care of older people, according to demographic prognoses, will become a serious problem around 2050 and thus might be an engine of economic growth, particularly in developed countries. First after a longer time of broad social use of such elements, social resistance might be lowered and some form of radical evolution might take place – but probably not before 2100.

This is mine reservation to the typical fantasy about radical evolution: they do not consider social forces and conflicts that will accompany biotechnological revolution. We can expect, however, a slow but dramatic change of social structure and inter-generational relations in the 21st Century as a result of informational revolution. Already in industrial civilization, but especially during informational revolution, together with the change of social role of women, a significant change of the social model of a family is taking place. The traditional model is a large number of children as an insurance for the old age of parents. Professional careers of women resulted in delaying the birth of children to an older age and generally in a smaller number of children per family. The insurance for the old age was expected in the form of a social insurance system; this system, however, in its classical form does not endure the growing share of old people in a society. Old people with a small number of children do not expect from them significant material help, they at most limit their respective help. Nevertheless, the return to the classical model of family is not possible: educated and professionally active women will not return to their classical roles. However, the collapse of traditional social insurance systems implies a new conflict: who should finance the living of old people? Therefore, financing the beginnings of biotechnological revolution is also questionable – who will finance it, if not old people?

One possible answer is so called *netocracy* (see [5]): only the new rich that will have both the financial and political power in the networked society will enjoy the possibilities of biotechnologies. However, I do not believe in the inevitability of destruction of democracy by informational revolution. Each revolution of such magnitude creates of course new social divides and new rich; but the industrial revolution did not destroy, only helped to create modern democracy, and the informational revolution has many aspects that support further development of democracy. The conflict about the property of knowledge will draw attention to the necessity of preserving democracy, as noted already by Thomas Jefferson (1813) [18]: a free access to ideas is both a necessary condition of democracy and helps to strengthen it. I do not believe, neither, that the new rich will so easily take the risks of testing new biotechnologies.

There are no definite answers to such questions. However, with respect to the new model of family, it is clear that a less inter-generational integrated model will gradually emerge, with lesser obligations between generations. This does not mean a clear cut of such obligations, but elder people will try to use diverse methods of increasing financial security, including a prolongation of professional activity, taking advantage of their life-long experience. Clearly, this also will provoke socio-economic conflicts: entrepreneurs might prefer employing only young people, older people might raise an issue of non-discrimination (not only because of race or sex, but also because of age) and fundamental rights of people.

Such problems might have also an impact on the reversal of the trend of urbanization of the world or the beginnings of *global forest*. Already today, the cost of living in villages and small towns is significantly smaller than in large agglomerations; an *escape to the forest* might be strongly motivated economically. In conclusion, the biotechnological revolution will be possible when humanity will overcome ecological threats and if this revolution will help in achieving a relative (obviously not absolute) ecological equilibrium.

2.4. Global Governance

The most important challenge, however, facing humanity before 2050, is in my opinion the issue of *creating new world order or global governance*. In fact, this is a direct consequence of information revolution and resulting globalization: people of the world perceive increasingly more their responsibility for global issues. Information technology and biotechnology will drive the changes of the world, responding to the broad social needs and demand. However, if we leave the satisfaction of this demand to big corporations (not to the free market which, as noted above, does not exists any more in its classical, ideal form on high technology markets) then we should expect next big crisis, a successive big bubble of artificially created demand motivated by the profits of oligopolistic market, not by solving the problems tormenting the world. Big corporations will of course do everything to hinder the emergence of a global governance – by promoting, e.g., the self-serving theses that the less government the better, that best brains require biggest rewards (and biggest compensations in case of failure), etc. But I hope that the lessons from the last crisis will show the emptiness of such neoliberal slogans and arguments.

The vision of the world governed by big corporations is not acceptable, leads to instability. As long as one big country – the USA – dominated the world order, dealing with the excesses of big corporations could be left to it, even if European Union and other big players often represented other interests. However, once a larger role in the world economy will be played by the most populous countries – China, India, Brasilia – a new world order and new institutions for global governance, a forum for achieving consensus between the biggest players will be needed.

There are many possible ways of creating new global governance, but two are the most probable. One is a renewal and strengthening of the role of United Nations Organization. If this organization will not address (because either of internal weakness or of lack of commitment of biggest players) new goals, entitlements and obligations, then another organization, perhaps between existing today, will have to fill out this void and help to create a new order. The tasks of such organization should be partly political, concerning global security (limiting armaments, eliminating military conflicts, alleviating diverse local and regional conflicts, etc.), partly economic regulatory (control of oligopolistic collusions and monopolistic aspirations of large corporations, regulation of international banking, etc.), partly globally developmental (supervision of global projects such as on the Antarctic or the Moon, other planets, etc.).

The creation of new global governance is a big challenge, larger than, e.g., emergence of European Union. But humanity must rise up to this challenge if we want to look to the future with trust.

3. The Challenges from and for Future Internet

From the above vision it is clear that the Future Internet will have fundamental impacts on the socio-economic development. The challenges *from* the Future Internet indicated above include:

– the conflict between global corporatization and global governance;

– the conflict about property of knowledge, thus concerning direct and indirect limits to freedom;

– the trend towards elitarism (or *netocracy*) inherent in current internet development dynamics versus democracy, with related issue of reform of education systems;

– the issue of ambient intelligence versus human rights;

– the issue of radical human evolution versus human sovereignty, with related threat of network and computer domination over people.

The main thesis of this paper is that the solution of and the challenges *for* the Future Internet should anticipate and take account of the predicted challenges *from* its future social application; the Future Internet should be based on a vision how to respond to such challenges. The existing Internet was also based on such a forward-looking vision: it was designed to warrant interconnectivity even under severe perturbations and to protect rights of every user of the network (even if these principles resulted from military considerations of warranting interconnections under nuclear attack and enabling the shift of command to anyplace). This was naturally achieved at the cost of disregarding its future commercial applications and has led to known dilemmas today (e.g., the tussle between peer-to-peer (P2P) and interactive network use, e.g., [19]). However, such design omissions can be solved technologically rather easily, good solution proposals already exist (e.g., the *trilogy architecture* [20]).

Quite different issue is a vision taking into account the challenges listed above. It is not sufficient to take into account some hypothetical scenarios, even if reasonable (see, e.g., [21]) without a consistent vision of the future world and a commitment what type of solutions should be preferred for this world.

Should Future Internet support oligopolization of the future integrated media, leading to direct or indirect censorship of ideas expressing the interests of big corporations, or should it rather promote free exchange of ideas and thus support global governance based on a direct opinion exchange? An answer saying that Future Internet solutions must be politically neutral is misleading: no technology is absolutely politically neutral, it can be used more or less easily in this or that political interest.

Therefore, we cannot dismiss questions such as:

- Should Future Internet help to exact strict intellectual rights and thus help in further privatization of knowledge, or should it rather promote open access to as much public knowledge as possible?

- Should Future Internet take digital divide and the trend towards netocracy as granted, or should it rather promote new forms of democracy and help to spread democracy by supporting educational reforms?

- Should Future Internet (or so unfortunately called real world Internet, see, e.g., [21]) be based on the assumption that ambient intelligence will be accepted by people, because market demand for it will be created by the advertisements expressing the interests of big corporations, or should it rather be based on the question which ambient intelligence applications are most likely to be socially accepted because they respond to true social needs without violating basic human rights?

- Should Future Internet be based on the assumption that a total immersion of a human being into a virtual world is desirable because anyway radical human evolution will occur, including such total immersions, or should it rather respect and support human sovereignty and dignity?

These are only questions, not answers, but I believe that it is not sufficient to limit socio-economic considerations of Future Internet to neoliberal convictions that market mechanisms would solve all problems. Therefore, such questions should be asked before constructing Future Internet.

4. Conclusions

It is difficult to summarize by classical conclusions a paper devoted mostly to a vision. Instead, I will try here to repeat here some of most important theses of this paper.

- The theories of a co-evolution of social attitudes and technological solutions are applicable only to a near-equilibrium evolution, certainly not applicable to such events as the beginnings of industrial revolution or beginnings of informational revolution, the emergence of Internet, or even the construction of a Future Internet. A technological solution of this magnitude of social impact was until now – and should be in future – based on a vision that hopefully would meet social expectations at least 40 years from its conception. Internet was based on such a vision, even if its success has outgrown the scope of that vision. Every radically new technology (computers, transistors, mobile telephony, digital television) is usually conceived at least 40 years before its broad social use and is based on a vision, not on a co-evolution of social attitudes and technological solutions.
- To construct Future Internet worth its name it is necessary to have a vision of the world in 2050. The paper presents such a vision – which is optimistic in the belief that humanity will be able to cope with most challenges and problems, conservative in the belief that social customs and attitudes do not change easily and require special reasons for change, finally radical in the belief that informational revolution already has had fundamental socio-economic impacts, between others invalidating most of neoclassical economics when applied to high technology markets.
- The solution of and the challenges *for* the Future Internet should anticipate and take account of the predicted challenges *from* its future social application; the Future Internet should be based on a vision how to respond to such challenges. This vision might be based on the vision of the world in 2050 discussed in most parts of this paper.
- The challenges *from* the Future Internet indicated above include:
  - the conflict between global corporatization and global governance;
  - the conflict about property of knowledge, thus concerning direct and indirect limits to freedom;
  - the trend towards elitarism (or *netocracy*) inherent in current Internet development dynamics versus democracy, with related issue of reform of education systems;
  - the issue of ambient intelligence versus human rights;
  - the issue of radical human evolution versus human sovereignty, with related threat of network and computer domination over people.

References

[1] S. Ratsanamy, S. Shenkers, and S. McCanne, "Towards an evolvable Internet architecture", in *Proceedings of 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*. New York: ACM, 2005, pp. 313–324.

[2] J. M. Rubina, "Challenges of Internet evolution: attitude and technology", in *Towards the Future Internet: A European Research Perspective*, G. Tselentis *et al.*, Eds. Amsterdam: IOS Press, 2009, pp. 12–23.

[3] A. P. Wierzbicki, "Delays in technology development: their impact on the issues of determinism, autonomy and controllability of technology", *J. Telecommun. Inform. Technol.*, no. 4, pp. 1–12, 2008.

[4] M. Castells, *End of Millenium: The Information Age*, vol. 1, 2, 3. Oxford: Blackwell, 2000.

[5] A. Bard and J. Söderqvist, *Netocracy, the New Power Elite and Life after Capitalism*. London: Pearson Education, 2002.

[6] A. P. Wierzbicki, "Megatrends of information civilization", in *The Knowledge-Based Economy: The Global Challenges of the 21st Century*, A. Kukliński, Ed. Warsaw: KBN, 2000, pp. 239–244.

[7] *Towards the Future Internet: A European Research Perspective*, G. Tselentis *et al.*, Eds. Amsterdam: IOS Press, 2009.

[8] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 2007.

[9] J. Garreau, *Radical Evolution*. New York: Doubleday, 2005.

[10] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens, *Limits to Growth*. New York: Potomac Associates, 1972.

[11] M. Heidegger, "Die Technik und die Kehre", in M. Heidegger, *Vorträge und Aufsätze*. Pfullingen: Günther Neske Verlag, 1954.

[12] A. P. Wierzbicki, "Technology and change", in *Creative Environments: Issues of Creativity Support for the Knowledge Civilization Age*, A. P. Wierzbicki and Y. Nakamori, Eds. Berlin-Heidelberg: Springer, 2007.

[13] N. Stehr, *Knowledge and Economic Conduct*. Toronto: University of Toronto Press, 2002.

[14] A. P. Wierzbicki and Y. Nakamori, *Creative Space: Models of Knowledge Creation Processes for the Knowledge Civilization Age*. Berlin-Heidelberg: Springer, 2006.

[15] *Creative Environments: Issues of Creativity Support for the Knowledge Civilization Age*, A. P. Wierzbicki and Y. Nakamori, Eds. Berlin-Heidelberg: Springer, 2007.

[16] J. Boyle, *The Public Domain*. New Haven: Yale University Press, 2008.

[17] A. P. Wierzbicki, A. Kameoka, and Y. Nakamori, "The new era of knowledge civilization and its episteme", in *Futurology – The Challenges of the XXI Century*, A. Kukliński and K. Pawłowski, Eds. Pruszków: Rewasz, 2008, pp. 45–82.

[18] *Thomas Jefferson Writings*. New York: The Library of America, 1984.

[19] D. Hausheer, P. Nikander, V. Fogliati, K. Wünstel, M. A. Callejo, S. R. Jorba, S. Spirou, L. Ladid, W. Kleinwächter, B. Stiller, M. Behrmann, M. Boniface, C. Courcoubetis, and M. S. Lee, "Future Internet socio-economics: challenges and perspectives", in *Towards the Future Internet: A European Research Perspective*, G. Tselentis *et al.*, Eds. Amsterdam: IOS Press, 2009, pp. 1–11.

[20] L. Burness, P. Eardley, and R. Hancock, "The trilogy architecture for Future Internet", in *Towards the Future Internet: A European Research Perspective*, G. Tselentis *et al.*, Eds. Amsterdam: IOS Press, 2009, pp. 79–90.

[21] F. Forest, O. Lavoisy, M. Eurich, J. Van Gurp, and D. Wilson, "Roadmap for a real world Internet applications – socioeconomic scenarios and design recommendations", in *Towards the Future Internet: A European Research Perspective*, G. Tselentis *et al.*, Eds. Amsterdam: IOS Press, 2009, pp. 325–334.

Andrzej Piotr Wierzbicki got his M.Sc. in 1960 in telecommunications and control engineering, Ph.D. in 1964 in nonlinear dynamics in control, and D.Sc. in 1968 in optimization and decision science. He worked as the Dean of the Faculty of Electronics, Warsaw University of Technology (WUT), Poland (1975–1978); the Chairman of Systems and Decision Sciences Program of International Institute for Applied Systems Analysis in Laxenburg n. Vienna, Austria (1979–1984). He was elected a member of the State Committee for Scientific Research of Republic of Poland and the Chairman of its Commission of Applied Research (1991–1994). He was the Director General of the National Institute of Telecommunications in Warsaw (1996–2004). He worked as a research Professor at Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, Japan (2004–2007). Beside teaching and lecturing for over 45 years and promoting over 100 master's theses and 20 doctoral dissertations at WUT, he also lectured at doctoral studies at many Polish and international universities. Professor Wierzbicki is an author of over 200 publications, including 14 books, over 80 articles in scientific journals, over 100 papers at conferences; the author of 3 patents granted and industrially applied. His current interests include vector optimization, multiple criteria and game theory approaches, negotiation and decision support, issues of information society and knowledge civilization, rational evolutionary theory of intuition, theories of knowledge creation and management.

e-mail: A.Wierzbicki@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Paper

Telecommunications, Universal Service and Poverty in Mexico: a Public Policy Assessment (1990–2008)

Cristina Casanueva-Reguart and Antonio Pita S.

Abstract—This article analyzes the design and implementation of telecommunications service policies targeted at the poorest regions of Mexico (1990–2008). It begins by defining universal access and service policies, their economic and social rationale. Secondly, it discusses the scope of public policies on universal service provision designed by Mexican authorities to achieve the goal of universal access. Thirdly, the paper analyzes the distributive effects of this set of policies among the poorest sectors of the population. The sources on which this research was based were two national surveys: the *Household Income and Expenditure Survey* (2008), and the *Household Survey of the Access and Use of Information Technologies* (2007). The additional information on regional economic development was based on the poverty indexes by the national population council and economic information given by Mexico's Census Bureau. Additional use was made of the annual reports prepared by Ministry of Communications, statistics published by the Federal Telecommunications Commission and official documents prepared by the government agencies. Finally, a series of in-depth interviews was conducted with the former representatives of the Office of Rural Telephony. Finally, the article discusses, in the light of available evidence, possible explanations for the apparent failure of the universal service policy that was implemented to bring at least basic voice services to Mexico's neediest.

Keywords—development, digital divide, market power, regulatory capture, social inclusion, universal service.

1. Introduction

The telecommunications services infrastructure is an important factor for economic development and social inclusion, and a crucial component leading to greater equality when services are available to and affordable to any person, irrespective of levels of income and geographic location.

In developing countries gaps remained in the market mainly because of regulatory failure, combined with exceptionally challenging geography and extremely low population densities, isolation and extreme poverty [1]. Government, institutions and public policy design have been faced with a situation where authorities had a social obligation to ensure that their people had access to basic telecommunications, but the ability to enforce these obligations entailed increasing difficulties. Some of these difficulties are related to the asymmetry between the government regulatory bodies and the power of the operators, particularly in the case of an incumbent operator with market power in the supply of almost every telecommunications service[1].

This article analyzes the design and implementation of telecommunications service policies targeted at the poorest regions of Mexico (1990-2008). It begins by defining universal access and service policies, as well as their economic and social rationale. The article then discusses the scope of public policies on universal service provision designed by Mexican authorities to achieve the goal of universal access, namely:

- Monitoring by the regulatory authorities of compliance on the part of the incumbent telecommunications operator, Teléfonos de México (Telmex), in fulfilling its social obligations as the dominant operator, following the company's privatisation (1990);
- Government policy aimed at providing connectivity – basic telephony services – to rural communities. This article presents the main research results, which focus on basic (voice) telecommunications services, both land-line and wireless[2]. This policy was deployed by private operators, mainly the incumbent. The participation of operators was based on a public tender processes organized by the government, which has resulted in an additional profitable business for the incumbent operator, with limited results in terms of access to these services by the poor.

Thirdly, the paper analyzes the distributive effects of this set of policies among the poorest sectors of the population. Finally, the article discusses, in the light of available evidence, possible explanations for the apparent failure of the universal service or social coverage policies that were implemented to bring at least basic voice services to Mexico's neediest, as well as the difficulties faced by the regulatory bodies behind the design and implementation of these policies.

[1] Almost 20 years after Telmex's privatization (1990) in 2009, Telmex's operated the 84.8% of the total the number of fixed line connections, and 72.3% of cellular lines. "The Competive Intelligence Unit" (CIU), http://octavioislas.wordpress.com/2009/07/23/3236-mexico-the-competitive-intelligence-unit-competencia-en-mexico-%C2%BFque-20-anos-no-es-nada-telecomunicaciones/

[2] The results and the discussion, presented in this article, are part of a larger project on Universal Access and Service which include telecommunications services and bandwidth in the provision of Internet, data and video transmission.

The sources on which this research was based were two national surveys: the *Household Income and Expenditure Survey* (2008), and the *Household Survey of the Access and Use of Information Technologies* (2007). The additional information on regional economic development reported in this article was based on the poverty indexes drawn up by the national population council and economic information given by Mexico's Census Bureau National Institute for Statistics, Geography and Information (INEGI), Bank of Economic Information (BIE), and housing and population counting, 2005. Additional use was made of the annual reports prepared by the Communications and Transport Secretary (Ministry of Communications), statistics published by the Federal Telecommunications Commission (COFETEL) and official documents prepared by the government agencies in charge of designing and monitoring the universal service telecommunications policies, principally the Ministry of Communications. Finally, a series of in-depth interviews was conducted with the former representatives of the Office of Rural Telephony, who were in charge of monitoring the implementation of social and universal telecommunications policies.

2. Access to Telecommunications Services as a Fundamental Right; Definitions of Universal Access and Universal Service: its Economic and Social Rationale

The recent literature on the universality of telecommunications services states, as a fundamental principle, that every citizen has the right to access telecommunications services with high quality standards. This implies fulfilling three basic conditions: universality, equality and continuity, defined as follows:

- universality: every citizen has the right to have access to telecommunications services with a high standard of quality;
- equality: universal access, irrespective of income levels and geographic location;
- continuity: ensuring that the service continues to maintain high quality standards, in an uninterrupted manner.

Regarding universality, the literature on this subject distinguishes between universal access and universal service:

- Universal access is when everyone has the right to use the service somewhere, in a public place. This service is also defined as communal or shared access. In general there would be at least one point of access per settlement over a certain population size, within a convenient and reasonable distance.
- Universal service is when every individual, household, business or institution can be provided with a service, using it privately, either at home or in an increasingly mobile variant, carried with the individual through wireless devices [2].

The rationale for a universal access or universal service policy is both economic and social. The economic rationale is based on the market's inability to provide infrastructure, connectivity and services on a universal basis. Thus, universal access and service policies are justified in the face of market gaps, so as to guarantee equality of economic opportunities, since telecommunications services are a critical component in the production of goods and services, as well as for social inclusion and, increasingly so, for political participation.

The social rationale consists of the will of the public policy makers who, as representatives of the state and its citizens, must guarantee social inclusion and avoid the exclusion of parts of the population, irrespective of income levels and geographic location.

Furthermore, the three defining characteristics of universal access and universal service are:

- availability: the service is accessible to inhabited parts of the country through public, community, shared or personal devices;
- accessibility: all citizens can use the service, regardless of location, gender, disabilities and other personal characteristics;
- affordability: the service is affordable to all citizens.

Thus, the policy of a universal telecommunications service consists of an explicit, direct and focused public policy, aimed at offering telecommunications services at prices that are affordable to the poorest sector of the population. The fulfilment of this policy requires a subsidy, since this sector of the population cannot afford these services at market prices. The subsidy may be applied on the supply or demand side of these services. On the supply side, it is usually applied through development and/or optimisation of the infrastructure (investment), which allows connectivity and thus takes into account the difficulty of infrastructure provision given the level of geographical dispersion of these communities, as well as the cost of providing the service.

On the demand side, the affordability of the provision of these services must be addressed, because the population currently without service generally lives in conditions of poverty, sometimes extreme, and thus their income does not allow them access to these services at the market price. Hence, when supplying these services is not profitable for the service operators, provision depends on various forms of subsidy [3].

3. Public Policies on Universal Service Provision in Mexico under the Responsibility of the Telecommunications Regulatory Bodies

According to the population counting of 2005, there were 184,748 rural communities in Mexico with a population of fewer than 2,500 inhabitants, and 197,479 communities of fewer than 5,000. These communities are inhabited by over 30 million men and women, which represented 29.1% of the Mexican population. In addition, a notable feature of these communities was their high level of dispersion, with 92.5% having fewer than 500 inhabitants.

Thus, in order to increase the social coverage of telecommunications services, a set of policies was drawn up and implemented. They represented the main public policies that have been put into practice to provide access and universal service in Mexico and among them, the most important were:

- Monitoring of Teléfonos de México's (Telmex, the incumbent operator) compliance with its obligations to provide a universal service, as set out in Telmex's licence of 1990, at the time that the public telecommunications operator was privatized. This licence was granted by the Ministry of Communications and Transport (Ministry of Communications[3]), which has acted as a regulator and has the authority to monitor Telmex's compliance with its obligations regarding universal service, rural telephony, as well as the modernization and expansion of the public network, contained in Telmex's licence.

- Rural telephony (1995–2007), aimed at offering services to communities with fewer than 500 inhabitants, with a direct subsidy from the Ministry of Communications.

- The creation of the *Social Coverage Fund* (Fondo de Cobertura Social, FONCOS: 2002–2007) by the Ministry of Communications aimed at increasing the coverage of rural telephony in communities with a population of between 400 and 2,499 inhabitants. For this purpose the regulator organized went out to public tender, calling for bids from telecommunications operators. The conditions of the tender included the provision of non-returnable monetary resources, which were originally allocated by the Ministry of Finance. In addition, the regulator allocated frequency bandwidth resources that were reserved for the purpose of social and universal service coverage, with a ten year licence (renewable) to use these frequency bands.

[3]In Mexico there are two main regulatory agencies directly involved with telecommunications, the Subsecretaría de Comunicaciones (Ministry of Communications), which is part of the Secretary of Communications and Transports (Ministry of Communications and Transport). The second regulatory agency is the Comisión Federal de Telecomunicaciones (Federal Telecommunications Commission), which is an autonomous government agency. In the case of universal service and universal access, the agency in charge of designing these policies is mainly the Ministry of Communications. This ministry has performed the major role in the design of universal telecommunications policy and in the surveillance of the implementation universal services by the operators.

3.1. Telmex Licence: Universal Service Obligations and Network Growth

The design of the universal service obligations included in the incumbent operator's licence should have addressed the main challenges posed by the gaps in the market, mainly in small rural communities generally located in remote and isolated areas and where the poorest people of the country live. This section deals with the context in which Telmex was privatized, which explains the scope of the clauses relating to universal service and rural telephony included in the licence. The scope of both the content and schedule for its implementation, as well as their impact on bringing connectivity to the neediest communities, was shaped to some extent by the context in which the privatization of Telmex took place. This section begins by describing this context, then briefly presents the content of the clauses relating to universal service obligations, namely, regarding network expansion, rural telephony and public telephones. It then presents the main results of an analysis of the impact of Telmex's fulfilment of these licence clauses.

In order to understand the government policy makers' limited leverage on the definition of the clauses relating to universal obligations or social coverage in Telmex's franchise agreement(1990), it is useful to analyze the specific juncture at which this process took place. During the 1980s, with the economy severely indebted, the burden of foreign debt and fiscal deficit had a major effect on the process of privatization of public companies and specifically on the approach adopted in the privatization of Telmex. Between 1965 and 1980 the economy had been growing at an average rate of 6.7%, but during the 80s this slowed to a yearly average of 1.8%. In this context, government policy makers embarked on an aggressive privatization program of public companies, with two purposes in mind: to increase the efficiency of Mexico's economy and to improve Mexico's public finances.

In the case of the public telephone company in Mexico, the potential revenue gain from this privatization and the public finance argument prevailed. The decision to privatize the public telecommunications operator in Mexico took place in 1989. During the privatization of Telmex, emphasis was placed on expected revenues. In fact, the privatization process was chaired by the Ministry of Finance and not by the Ministry of Communications and Transport [4]–[6].

To maximize revenues from privatization, the government sold to a single set of investors a package that included Telmex and Telnor, in addition to the only nationwide mobile network franchise, as mentioned above, the Federal

Microwave Network, as well as an ample bandwidth allocation. Thus, overnight, the emerging company became a formidable player in the sector. It was allowed to offer all types of telecommunications services with the exception of television broadcasting services. Thus, Telmex became a horizontally and vertically integrated telecommunications service provider with a nationwide network for all its services. Furthermore, it was guaranteed little or no competition in key services for several years. The new private owners of Telmex were given a *de jure* monopoly over the long distance markets (national and international) for six years. They inherited the monopoly over local telephone services. By creating a horizontally and vertically integrated telecommunications company, the government could receive a higher price for privatizing the firm and reach a short-term public finance goal. For potential buyers, the company was very attractive [4], [5]. They were allowed to buy a stream of excess profits sustained by a monopoly, more valuable than a stream of revenue generated under competitive conditions [7].

In the negotiation process that accompanied the privatization of Telmex, government policy makers let the collection of revenues from privatization preside over other goals such as economic efficiency, well-being and social inclusion. This context explains the lack of leverage or bargaining power of the authorities, specifically on the subject of social coverage. Although Telmex's license included clauses governing universal service obligations, rural telephony and public telephone booths, as well as network expansion, Telmex's commitment to these clauses ended in 1994 and the results lagged very much behind the objective of providing a basic universal telecommunications service to rural areas.

The following paragraphs give a summary of the main clauses relating to Telmex's universal service and network growth obligations. The result of an empirical analysis is also presented, showing the outcomes of the implementation of these policies, reflecting the achievements of network growth, rural telephony and public telephone services or telephone booths (1990–1998).

From a regulatory perspective, Telmex's licence made this company operate as a regulated monopoly. The company was given a set of operational goals that it was required to meet:

- to expand the number of basic telephone lines by a minimum of 12% per annum, until the end of 1994;
- to continually reduce the waiting period for the basic telephone service in localities with automatic switching capabilities to a maximum of one month by the year 2000;
- to provide a public payphone service to every locality with more than 500 inhabitants by the end of 1994, and to increase the penetration of public telephone booths from 0.5 per thousand inhabitants, to five per thousand by the end of 1998;
- to provide a public payphone service to every locality with more than 2,500 inhabitants (less than 5,000 inhabitants, according to the definition of rural community), if there were at least 100 applications from potential end users, and an up front payment equivalent to three months' line rental. "*After these conditions were fulfilled (...) Telmex would deliver the service within a time frame of not more than 18 months*"[4];
- to publish a four year (network) expansion and modernization programme, in accordance with the goals set by the licence, and to agree with the Ministry of Communications on programmes for rural telephony and public telephone booths.

3.1.1. To Expand the Number of Basic Telephone Lines by a Minimum of 12% per Annum

The 12% telephone line expansion requirement ended only four years after the 1994 privatization of the sector. Quantitative goals for a longer period would have conflicted with the goal of revenue maximization at the time of privatization. The results shown in Table 1 tend to support the view that the quantitative line expansion requirement set out in the concession or Telmex licence was an effective regulation. In fact, during the years 1991 to 1994, Telmex's average annual line expansion was 11.8%, and thus close to meeting the 12% requirement. However, once such an expansion requirement ceased to exist, average line expansion fell to only 6.8% in the period 1994–2000 (see Table 1).

Table 1
Five year average growth in telephone lines and GDP (1965–2000)

Period	Telephone lines [%]	GDP [%]
1965–1970	12.8	6.9
1970–1975	12.9	6.5
1975–1980	10.5	6.7
1980–1985	6.4	1.9
1985–1990	7.6	1.7
1990–1994	11.8	3.6
1994–2000	6.8	3.5
Source: SCT, Anuarios Estadísticos (1965–2000).		

3.1.2. Rural Telephony: Basic Telephone Service to Communities with More than 500 Inhabitants (1990–1998)

As a result of the negotiations between government policy makers in the field of telecommunications and the group of investors, Telmex's licence freed them from their obligation to serve communities with fewer than 500 in-

[4]SCT, Modificación a Teléfonos de México, 1990, http://www.cft.gob.mx/work/sites/Cofetel_2008/resources/LocalContent/3964/1/10ago90.pdf

Table 2
Telephone service to communities with more than 500 inhabitants

State	Lines per 1,000 inhabitants towns 500 to 2,499	Lines per 1,000 inhabitants towns 500–4,999	Rural communities communicated by Telmex, 1990–1994	Total population towns 500–2,499	Total population towns 500 to 4,999	GDP per capita
Chiapas	1.00	0.44	950	951,521	2,136,825	37.8
Oaxaca	1.24	0.65	1,362	1,095,547	2,102,278	39.8
Tabasco	1.31	0.78	678	517,227	863,855	47.6
Guerrero	1.27	0.63	932	732,388	1,470,855	47.8
Tlaxcala	0.77	0.37	109	141396	294,861	47.9
National	**1.35**	**0.65**	**16,738**	**13,339,307**	**27,937,529**	**83.2**
Campeche	1.43	0.70	135	94653	193,781	121.7
Quintana Roo	1.54	0.74	122	79123	164,691	126.4
Coahuila	1.65	0.70	238	144448	337,934	129.8
Nuevo León	3.30	0.79	232	70211	293,812	173.5
Distrito Federal	0.00	0.00	0	13268	26,550	188.0
Source: SCT, Annual Reports, INEGI, Census 1990.						

habitants, which according to the census of 1990 represented 21.16 million people or 47.2% of the inhabitants in rural communities in Mexico [8].

An analysis of the impact of Telmex's rural telephony operations points to very limited results. The impact on telephone density, following Telmex's compliance with requirements on basic telephone service provision in rural towns, was extremely low, even when telephone density was estimated as the number of lines per thousand inhabitants[5]. This estimate shows that the country's average telephone density was 1.35 lines per 1,000 inhabitants in rural communities (500 to 2,499 inhabitants), and if the definition of rural communities included "enlarged rural communities" (from 500 to 4,999 inhabitants), the telephone density estimate drops to half the previous figure, or 0.65 lines per thousand inhabitants. Based on this analysis it is possible to assert that fulfilment of overall requirements, provision of rural telephony and the installation of public telephone booths in rural areas, had very much fallen behind in relation to the goals set by Telmex's licence (see Table 2, which presents the telephone density, according to our definition, in the five more prosperous states and in the five poorest states in Mexico).

Thus, in spite of the fact that public telephone booths were the strategy mostly used by Telmex to fulfil its universal or social obligations, compliance with the commitment of providing public access through public telephone booths was insufficient. At the end of 1998, Telmex admitted that it had only installed 3.19 public booths per 1,000 inhabitants. Taking into account that Mexico's population at the time was 96 million, Telmex would have had to install at least 480 thousand public booths in order to comply with the social obligation dictated by its licence. Unfortunately,

[5]See Caslon analytics, metrics and statistics, http://www.caslon.com.au/metricsguide8.htm

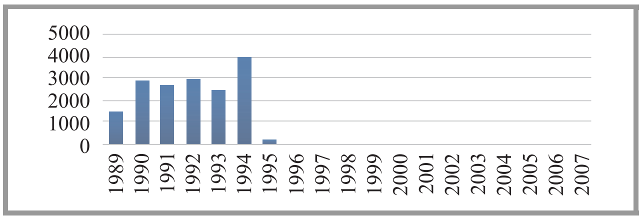


Fig. 1. Number of rural communities with more than 500 inhabitants, communicated by Telmex with at least one telephone line (1990–2007).

according to the definition of universal service set out in Telmex's licence and because of the government's lack of leverage at the time of privatization, the universal service obligation ceased and areas that were served with at least one public booth increased slightly in 1995 and 1996 and ceased to grow indefinitely from 1997 (see Figure 1 and Table 3).

Table 3
Telephone lines provided by Telmex in communities with a population of 500 or more inhabitants

Acumulated	Annual [number]	Annual [%]
4,350	2,854	190.8
16,542	4,006	32.0
Average annual growth: 1990–1994		**93.4253**
16,735	193	1.2
16,738	0	0.0
16,738	0	0.0
16,738	0	0.0
Average annual growth: 1994–2007		**0.0014**
Sources: SCT, Annual Reports (2000 and 2007).		

Table 4
Teledensity and level of development in different states in Mexico, 2008

State	Lines per households [%]	Non residential lines[1] [%]	Mobile lines[2]	GDP per capita	Poverty index
National	52.2	12.7	71	96.8	
Chiapas	19.1	4.2	41.1	37.8	Very high
Oaxaca	26.4	4.6	39.8	39.8	Very high
Tabasco	28.1	6.7	75	47.6	High
Guerrero	40.6	7	44.8	47.8	Very high
Tlaxcala	40.3	4.4	44.1	47.9	Medium
Campeche	32.7	6.4	71	121.7	High
Quintana Roo	38.0	16.1	95.9	126.4	Low
Coahuila	63.3	12	86.6	129.8	Low
Nuevo León	79.9	22.2	91.8	173.5	Very low
Distrito Federal	104.6	40.5	102.7	188	Very low

[1] lines per one thousand employed people, [2] lines per one hundred people.

Sources: Cofetel, 2008; INEGI, 2005 and 2008, National Population Council, 2005.

According to Telmex's license, the commitment to provide a basic service under the premise of universal access through public booths in Mexico is far outweighed by the challenge of providing services to the poorest communities of Mexico. This is true more particularly in the context of a large deficit of telecommunications services in the country as a whole: on average, there is provision to only 5 households out of every ten (52.2%) and 12.7 non residential lines per one thousand employed people (see Table 4).

Table 5
Average revenue per capita in selected developing countries, 2007

Latin America	Country	ARPU	GDP per capita	ARPU/GDP per capita[1]
	Colombia	131.0	7,400	1.8
	Mexico	178.0	12,400	1.4
	Brazil	17.2	9,500	0.2
	Chile	17.0	14,300	0.1
	Argentina	12.2	13,100	0.1
Emerging Europe	Hungary	26.5	19,300	0.1
	Czech R.	28.5	24,500	0.1
	Poland	17.0	16,200	0.1
	Ukraine	7.0	7,000	0.1
	Russia	10.0	14,800	0.1
Africa/ Middle East	Turkey	13.7	12,000.00	0.1
	Iraq	12.3	3,700.00	0.3
	South Africa	20.7	9,700.00	0.2
	Egypt	9.7	5,000.00	0.2
Emerging Asia	India	8.6	2,600	0.3
	China	10.8	5,400	0.2
	Korea	45.0	25,000	0.2
	Taiwan	23.0	30,100	0.1
	Singapore	348.0	49,900	0.7
	Hong Kong	22.5	42,000	0.1

[1] ARPU – average revenue per user: the estimation gives ARPU as a percentage of GDP per capita, in each country.

Sources: Merrill Lynch, 2008, CIA World Factbook, 2008, and Office of Rurar Telephony, 2009.

Most likely the deficit of connectivity has been compensated for by the use of cellular or mobile services, at market prices. A comparative analysis of revenue per minute from wireless services suggests that Mexico has one of the highest tariffs for mobile services (see Table 5).

The following section presents information relating to the "agreements" between Telmex and the Ministry of Communications, which according to its licence, were supposed to continue after 1994.

3.1.3. Telmex Agreements with the Ministry of Communications on Programmes for Network Expansion, Rural Telephony and Public Telephone Booths (1995 to date)

In 1995, according to Telmex's licence, the company was required every four years to establish a programme for network expansion and provision of rural telephony and public telephone booths. However, in spite of the fact that an exchange of documents took place between Telmex and the Ministry of Communications in 1995, this exchange did not materialize into an action programme.It was not until 1998 that this took place, when Telmex set up a rural telephony programme, supposedly in an agreement with the Ministry of Communications, retroactive to 1995 (1995–1998).

In December 1998 Telmex sent a report on the fulfilment of the goals of this programme. The main goal achieved was the additional coverage, compared to 1994, of 4,288 communities, through public telephone lines or public booths.

However, strictly speaking, the service provided to these additional communities was part of a rural telephony project undertaken by the Ministry of Communications. Telmex was involved as the winner of a public tender in which they were chosen as the provider, but the project was financed by the Ministry of Communications (the results are presented below of the universal service policy undertaken by the Ministry of Communications).

It was only in July 2006, when Telmex delivered its results on network expansion and rural telephony for the periods 1995–1998, 1999–2002 and 2003–2006. The documents handed over by Telmex to the Ministry of Communications were accompanied by an appeal for confidentiality and the Ministry of Communications accepted the condition of confidentiality requested by Telmex. Nevertheless, the annual reports published by the Ministry of Communications (2001–2006) showed that since 1994, Telmex made a very small contribution in this area.

The privatization of Telmex and the attempts to regulate the monopoly through this company's license did not bring a strengthening of the regulatory authority's monitoring and enforcement capabilities. This had a severe effect on achieving the goals set by the universal service or universal obligations defined in Telmex's licence following privatization. Studies on telecommunications reform suggest that privatization by itself, without a strong regulator, does not yield significant performance improvements in the telecommunications sector [5], [7], [9].

4. Rural Telecommunications Services for Communities of Fewer than 500 Inhabitants: Direct Government Subsidy

This section analyzes the policy directly implemented by the Ministry of Communications, aimed at providing telecommunications services to rural communities of between 100 and 499 inhabitants. These programmes originally focused on small towns and villages with fewer than 500 inhabitants (1990–2002), but later on, with the establishment of the *Social Coverage Fund* (FONCOS), the focus of these programmes shifted to communities of between 400 and 2,500 inhabitants.

Table 6
Rural telephony supplier

Cellular company	Lines	%
Telcel	8,358	25.1
Iusacell	11,012	33.1
Telecomm	13,772	41.4
Others	100	0.3
Total	**33,242**	**100.0**
Source: Office of Rural Telephony.		

The programme was directly financed by the Ministry of Communication (1995–2006) and was aimed at communities with fewer than 500 inhabitants. It involved the main providers of mobile telecommunications services, including Telcel, the mobile company part of the same group as Telmex (25%), and the public satellite company, Telecomm (41.4%, see Table 6).

The size of these targeted communities confirmed that the policy successfully focussed on the poorest towns of Mexico and even in cases where regions of higher income were served, the subsidy focussed on the neediest rural communities that are generally located in remote and isolated areas (see Table 3). Nevertheless, telephone density in these small towns remained extremely low, as was the case of the services provided by Telmex in larger communities, where the estimated telephone density is 0.45 lines per 100 inhabitants when population data for these communities is taken from the 2000 census and 0.44 with population data is from 2005 (counting of population and housing, 2005, see Table 7). It is also likely that the actual telephone density in these small towns was higher because of the use of mobile technology, when it was available in these regions.

An additional source of information consisted of in-depth interviews with the former representatives of the Office of Rural Telephony, part of the aforementioned Ministry of Communications, where data compiled by this office is based on their fieldwork aimed at verifying the correct operation of the installed lines. This information complements that of the previous analysis[6].

The information provided by the Office of Rural Telephony showed that 33,242 lines were installed between 1995 and 2006. An analysis of this information also showed a very rapid growth in the number of installed lines between 1995 and 2000 (135.42% yearly average growth) and that the pace of growth declined considerably over the following years, where the yearly average growth observed between 2001 and 2005 was only 1.16%. There was no evidence of growth between 2006 and 2009.

The results of the former analysis are even more dramatic considering the outcome of the fieldwork aimed at verifying the correction operation of the equipment. Here the data showed that only 41.5% of the installed lines were in operation and out of these, 58.5% of the lines were out of service and abandoned. The interviews with the former representatives of this office suggested that the people of these communities were gradually shifting to mobile technology, where this service was available, in spite of the higher costs involved in the use of mobile communication.

The former observations raise doubts over the commitment made by the government for bridging the connectivity gap in the smallest and poorest communities of Mexico. It also

[6]Is worth mentioning that there is an inconsistency between the information accounted by the Annual Report (2007) and the Office of Rural Communication, the first source reported 34,676 installed lines, an the Office of Rural Telephony accounted for 33,242, the difference between the two sources is 1,434 installed lines.

Table 7
Rural telephony, lines installed by Secretary of Communications, towns between 100 and 499 inhabitants, 1995–2007

	Communities	Lines 2007	Teledensity /100, 2000	Teledensity /100, 2005	GDP per capita
Total	**184,748**	**34,676**	**0.45**	**0.44**	**70.88**
Chiapas	19,237	3,560	0.48	0.42	28.6
Oaxaca	10,025	2,540	0.41	0.37	32.5
Tlaxcala	1,138	117	0.41	0.41	37.3
Michoacán	8,965	1,861	0.45	0.45	39.9
Chihuahua	12,095	896	0.43	0.53	102.9
Quintana Roo	1,800	177	0.46	0.44	107.5
Campeche	2,595	240	0.49	0.35	121.7
Baja California	3,918	248	0.56	0.56	93.0
Campeche	2,595	240	0.49	0.35	121.7
Nuevo León	5,169	561	0.50	0.57	133.1

Source: INEGI, Censo de Población y Vivienda, 2000; Conteo Población y Vivienda, 2005; Ministry of Communications, *Annual Report* (2000–2007) and Bank of Economic Information (BIE), 2007.

raises questions over the nature of the agreements signed by the Ministry of Communications and the operators undertaking the installation of the telephone lines, specifically operation and maintenance in accordance with acceptable quality standards. The personnel interviewed agreed that the contracts included maintenance and quality clauses, which pose additional questions on the strength of the ministry as a regulator capable of enforcing these clauses.

Table 8
Number of new telephones installed in communities with less than 500 inhabitants 1995–2008

1995	4,000
1996	9,369
1997	10,545
1998	20,208
1999	23,063
2000	31,083
Average annual growth 1995–2000	**135.42**
2001	31,083
2002	31,453
2003	31,820
2004	32,326
2005	32,841
2006	33,240
2007	33,242
Average annual growth 2001–2008	**1.16**

Source: Ministry of Communications, Office of Rural Telephony.

5. Social Coverage Fund (FONCOS)

In 2002 the *Social Coverage Fund* (FONCOS) was established as a trust fund with an allocation of 75 million US dollars[7] provided by the Ministry of Finance to the Ministry of Communications. Its main purpose was for the funding of social telecommunications services, focussed on serving communities of between 400 to 2,500 inhabitants. The Ministry of Communications designed two different public tender processes. The first, *Basic Telephony 1* (STB-1 to use its Spanish acronym), was aimed at installing public telephone lines in communities of extreme poverty. The second, *Basic Telephony 2* or STB-2, focussed on communities with higher levels of income. For STB-1, the subsidy for the chosen operator consisted of both financial and bandwidth resources for 10 years (renewable), which were reserved by the government for social coverage purposes. The subsidy to the end user included all expenses relating to the installation and rental of the equipment, so that the end user had only to pay for call traffic, charged for via prepaid cards.

For STB-2, the subsidy to the winning operator consisted of bandwidth resources only. The end user was charged for installation costs and call traffic, exonerating them from payment for the rented equipment. In this case, the subsidy for the chosen company consisted only of the licence to operate bandwidth resources for 10 years (also renewable). Although in the first round four companies participated, two of them were Telmex and Telcel. The latter is a mobile service provider belonging to the same consortium as Telmex. In the second round of the tender process, Telmex was the only bidder. In both public tenders Telmex, the incumbent operator, was chosen. The contract with this incumbent operator was signed on February 2005 with the target of serving 5,979 communities.

There were two changes to the terms of the contact signed between Telmex and the Ministry of Communications. The first was related to the inability to serve 737 communities

[7]The total sum accounted by 750 million pesos, the exchange rate between Mexican pesos and US dollars at the time was around 10 Mexican pesos per US dollar.

due to the fact that these towns lacked an electricity infrastructure or due to difficulties imposed by weather contingencies. The settlement consisted of a time extension granted to Telmex in order to serve 506 communities.

The second change to the original contract consisted of exchanging bandwidth resources reserved by the government for social coverage purposes for bandwidth with high commercial value for Telmex. This change had severe implications for both the implementation of the universal service process and in terms of the dominant control of infrastructure on the part of the incumbent operator. This latter implication had negative consequences due to the lack of competition in the telecommunications services markets, thus affecting the economy and society as whole.

In November 2006, a few weeks before the end of the presidential and ministerial administration of 2000–2006, an exchange of frequency bands took place: its 21 MHz allocation in the 1.5 GHz band, which was originally allocated by the Ministry of Communications to Telmex as part of the Social Coverage Fund, was exchanged for 10 MHz in the 450 MHz band. The Ministry of Communications did not exercise its power to monitor the use of these frequency bands.

The exchange of frequency bands turned out to be commercially convenient for Telmex, since the 450 MHz band was the most appropriate for the provision of wireless services with technology known as CDMA 450. Among the advantages of the use of frequency resources with this technology are:

- The ability to digitalize and interleave calls with a code attached to each one, allowing a large number of simultaneous calls without interference.
- An additional advantage consisted of having a larger coverage per cell, which requires a smaller number of cells, resulting in a more cost-effective technology. Also, the possibility of supplying a wide variety of services, such as Internet, telephony, data transmission, videoconferencing and connectivity between local networks.
- This frequency band also makes use of CDMA 2000 1X and CDMA 2000 1xEV-DO technologies, which allow for high speed data transmission, equivalent to the digital service line (DSL).

The exchange of bandwidth resources dedicated to social telephony for resources with ten years of high commercial value was carried out by the Ministry of Communications and allowed Telmex access to and use of these resources without going through an open public tender. This raised questions over Telmex's interest in participating in the social coverage tender process.

Former representatives of the Office of Rural Telephony argued that Telmex's true interest was to acquire the use of the frequency bandwidth resources, with a potentially high financial return, thus evading the higher transactional and monetary costs involved in taking part in an open public tender, which has been the allocation mechanism for radio bandwidth resources for commercial use drawn up by the government in accordance with the federal law on telecommunications (1995)[8].

The former analysis leads us to consider the role of the government authorities in organizing tender processes and allocating public financial and bandwidth resources for social coverage. In this case, the Ministry of Communications played a different role by granting valuable infrastructure resources to be used commercially, at a very low cost for the incumbent operator.

Additionally, and based on fieldwork and remote monitoring performed by the Office of Rural Telephony, the supervision of the services offered by Telmex under the *Social Coverage Fund* (FONCOS) showed that, out of the programme objective of 109,016 telephone lines (75,797 lines under the STB-1 program and 33,219 under the STB-2 program), only 88,791 were actually installed, which implies that 20,225 lines were never installed.

There was a brief period, after 2006, when the new administration of the Ministry of Communications verified the services delivered by Telmex, under the *Social Coverage Fund*. The Ministry Office of Rural Telephony identified numerous irregularities, for example the installation of two land line connections in the same household, which proved less costly for Telmex (19,397 lines). A similar discovery was made of lines that were not connected to any specific household, which obstructed the verification of their operation (6,983 lines). In contrast, before 2006, the Ministry of Communications had punctually paid Telmex, based on the invoices that the company presented. During a brief period of time the Ministry of Communications initiated a process to impose sanctions on Telmex and to suspend payments to the company. However, this process never transcended the boundaries of the ministry because different groups within the ministry restricted the sanctioning process. Furthermore, the group that initiated this process no longer serves in the Ministry of Communications.

Here again, the analysis reveals the role of the regulator, firstly in the tender process, and specifically, in the process of allocating bandwidth resources with a high potential return for Telmex. Furthermore, the regulator did not supervise the use of these resources, which were specifically allocated for social communications coverage. This finding suggests regulatory capture and corruption on the part of the regulator [10], taken to a serious extreme since the regulator did not exercise its power in preventing the re-allocation of resources originally targeted at the poorest people, which in turn strengthened the market power of Telmex. Secondly, the regulator did not impose sanctions on Telmex for its breach of the agreement on social coverage. The role of the regulator was eclipsed, most probably by numerous instances of lobbying, resulting in a failure to

[8] Federal Telecommunications Law: Article 14. The licenses for the use of radiofrequency bands for determined uses will be granted through an open public auction. The Federal Government has the right to receive the agreed monetary resources; http://www.diputados.gob.mx/LeyesBiblio/pdf/118.pdf

Table 9
Telecommunication services distribution according to households' income (deciles), 2008

Service	Deciles									
	1	2	3	4	5	6	7	8	9	10
Line connection	22.6	45.8	56.4	66.5	76.2	78.5	83.4	87.7	91.7	92.6
Mobile service[1]	22.1	42.2	52.0	65.2	70.4	77.7	82.1	87.0	91.1	86.7
Cable or satellite TV	5.0	12.7	17.5	26.2	32.6	41.5	50.3	62.3	72.3	75.8
Internet	0.2	1.7	3.4	7.2	11.4	17.4	27.2	41.4	56.1	60.1

[1] Mobile services are accounted when at least one member of the household has a mobile line.
Source: INEGI: ENIGH, 2008.

consider the well being and social inclusion of the poorest sector of the population.

So far the limited achievements of the different public policies aimed at providing universal service have been presented as being due firstly to the limited implementation of the clauses set out in Telmex's licence and secondly to the violation of various agreements, including the FONCOS contract with Telmex.

There are then at least three main findings that can be drawn from the previous analysis: firstly, that the provision of universal service or universal access has been extremely limited in addressing the market gaps in Mexico's rural areas and telephone density in the different services continues to be very low. Twenty years after privatization of the public telephone company and 15 years after the liberalization of the telecommunications markets in Mexico, connectivity and telephone density remain a major challenge for public policy in Mexico. The second major finding is the continual breaching by Telmex of its universal access or service commitments, not only as was originally stated as part of its licence, but later as the result of a contract that was signed with the Ministry of Communications making it the main supplier of these services. Finally, the third finding is the limited leverage of the telecommunications authorities and their difficulty in enforcing contracts and agreements, as well as in imposing sanctions.

There is an extremely low density telecommunications infrastructure in rural areas, with the exception of the mobile infrastructure existing in some. This has led end users to rely increasingly on mobile services, which are more costly than regular services. The following section analyzes the distributive effects of this set of policies among the poorest sectors of the population.

6. Access and Expenditure in Telecommunications Services by the Poorest Sectors of the Population and their Relation to Income Distribution

In this section, we analyze the distribution of telecommunications services as a function of different levels of household income. The starting point for this analysis is the decile distribution of households by level of income and their expenditure on telecommunications services as a proportion of their income. The source of information is the *Household Income and Expenditure Survey* for 2008, based on a nationwide representative sample. Each decile comprises the same number of households, which are ranked from the lowest to highest income. Comparing the lowest income decile with the highest, the latter is 30 times higher.

6.1. Access to Services

Those households within the lowest income decile (the poorest) have a significantly lower level of access to telecommunications services compared with households with higher income. This finding is consistent with the results previously presented on universal service provision and on the low density of infrastructure presented before. In the lowest decile, only two households in every 10 have a home telephone line connection, while in the highest income decile, nine out of every 10 have a land line connection in their homes. Very similar figures can be found for mobile services (see Table 9).

In the case of cable or satellite TV services, which have the technical capability for supporting telecommunications services and are currently used by many countries for that purpose, including urban areas of Mexico, the distribution of these services is even more skewed than land line connections or mobile services. Thus, in the lowest decile only 5% of households have access to TV based on this infrastructure, while 75% of households in the highest income decile have access to this service. The most dramatic case of this unequal distribution is found in relation to Internet access, where 60.1% of households in the highest income decile have access to the Internet, while in the lowest, only 0.02% have this service in their homes (see Table 9).

6.2. Expenditure on Telecommunications Services

The expenditure of the poorest households on telecommunications services as a percentage of their total outgoings is twice as high as the expenditure of the wealthiest households: 4.2% in the lowest deciles and 1.9% in the highest deciles. These results suggest that the demand for telecommunications services tends to be inelastic, which means that people tend to demand and use these services regardless of their income. The larger proportion of the expenditure of poorest families is explained by the fact that they live in remote and isolated areas and depend more on public telephone booths and mobile services, which tend to be more expensive (see Table 10). As mentioned before, mobile services in Mexico have one of the highest prices compared to other developing countries (see Table 5).

Table 10
Expenditure in telecommunications as a percentage of total expenses by decil (2006=100)

Decil	Average households income per quarter MEX pesos	Expenditure in telecommunications as a percentage of total expenses
1	3,320	4.2
2	7,174	4.1
3	10,042	4.3
4	12,739	4.3
5	15,845	4.4
6	19,506	4.5
7	24,246	4.2
8	31,472	3.8
9	43,796	3.2
10	99,215	1.9
Source: INEGI: ENIGH, 2008.		

An additional factor that induces greater expenditure for the poorest areas is related to the outdated definition of the areas of local service, which has no technical or economic (cost-related) basis and artificially classifies a call as long distance, incurring higher charge. This particularly affects people in rural areas, where the largest share of their traffic consists of long distance calls. The higher expenditure on telecommunications services has an impact on the opportunity for the poorest sectors of the population to spend in other areas like health, nutrition, education, home maintenance, or make a higher investment in productive activities.
These results also suggest that the provision of telecommunications services under the aegis of "universal access" or "universal service" is lagging behind the unfulfilled demand for these services. This leaves the poorest sectors of the population dependent on the supply of services at market prices, and mostly wireless services which are more expensive and difficult for them to afford.

7. Conclusions

Almost 20 years after Mexico's privatization of the telecommunications services by the incumbent operator, Telmex (Teléfonos de México), the premise of universal service is far from being fulfilled. Thus on average, only five out of every 10 households have access to a basic telephone service and in some states such as Tabasco and Chiapas only three, while in Oaxaca only two (1.9) out of 10 households have access to a telephone line. Similarly, when analyzing expenditure on telecommunications services, the collected data has demonstrated that in contrast to what has been the goal of the universal service policy in Mexico, namely social inclusion and overcoming poverty, the telecommunications policies have become a regressive tax for Mexico's poorest. These dramatic results are in direct contradiction with the fact that Telmex cheaply acquired radio bandwidth for social services by continually winning their public tenders, because regrettably, they later used them exclusively for commercial purposes.
In explaining the reasons for these poor results, the paper has also pointed out the shortcomings of the regulation in place and of the implementation of this regulation by the authorities, who are directly responsible for the failure to comply with the telecommunications service policies originally targeted at the poorest regions of Mexico. These shortcomings consist essentially of:

– restricting the application of services to communities with more than 500 inhabitants;

– allowing Telmex to choose between serving rural communities either by a public telephone booth or by a land line connection to households;

– allowing a policy far below the standard provided by the International Telecommunication Union;

– providing insufficient direct government subsidies for rural telecommunications services for communities with fewer than 500 inhabitants located in the poorest states in Mexico, who not only lack a basic telecommunications infrastructure but also present the lowest telephone density, seriously jeopardizing their right to use adult distant learning education;

– imposing strong limits on competition by establishing high barriers of entry for competitors in the mobile services market, which has allowed very high prices to be set for the end user and some of the highest among developing countries (as shown in Table 5 above).

These shortcomings show that universal access to telecommunications services in Mexico is a representative case of "regulatory capture" [10], [11], where the regulators and government authorities have been "captured" by the incumbent operator and have subordinated their regulatory power to a monopoly-based profit-seeking behavior.
Finally, although it is impossible to go back to the time of privatization, there is a long list of different regulations that have already been implemented by both developed and developing countries, which were created before the described policies for efficient universal access and service were imposed upon the incumbent operator's licence permit. These policies have achieved a larger density and better distribu-

tion of services and some of the following regulations could thus now be applied with greater chances of success:

- An obligation to provide interconnection on a non-discriminatory basis, according to high quality standards and establishing charge estimates on the basis of long term incremental cost. This applies particularly to those networks with market power: Telmex and Telcel.
- Unbundling the local loop, thus allowing non-discriminatory access to sections of the incumbent operator's network infrastructure.
- To guarantee a free flow of information on network capabilities, specifically on the points of presence and network architecture. This also applies particularly to Telmex and Telcel. This will contribute to creating incentives for new investment and the participation of new players, increasing coverage in regions previously lacking provision.
- To provide services across networks on a non-discriminatory basis, to high quality standards. It is known that international networks get better roaming services from Mexican networks than the cross-network services that national networks get between one another. This has been a barrier to entry for new players that have prevented investment and coverage in regions lacking coverage. It is worth reiterating that Telcel is the operator holding control of 75% of lines and operating the largest mobile infrastructure.
- To allow mobile virtual network operators (MVNO), which enable new players in the market to provide mobile phone services without necessarily having their own licensed bandwidth allocation, nor does it necessarily require them to have the entire infrastructure required to provide mobile telephone services.
- Re-defining the domain of local services, whose definition currently incurs an artificial increase in prices for so-called long distance calls. Without a technical basis for such a definition, this particularly affects those rural areas whose traffic is mainly long-distance based.
- Enforcing the declaration of Telmex as an operator with (monopolistic) market power and imposing upon it special requirements regarding quality, prices and information, so as to level the playing field by allowing other operators to enter the market and promote healthy competition.
- Closely monitoring Telmex and Telcel, in order to guarantee the proper delivery of telecommunications services to the poorer areas of Mexico.
- To ensure accountability and transparency in all legal processes relating to telecommunications services, regulation and competition. This not only provides legal certainty, but is a potential antidote for regulatory capture and corruption.

References

[1] S. Maddens, *Trends in Universal Access and Service Policies: Changing Policies to Accommodate Competition and Convergence*, in "Support for the Establishment of Harmonized Policies for the ICT Market in the ACP", Geneva, ITU, 2005 [Online]. Available: http://www.itu.int/ITU-D/treg/Events/Seminars/GSR/GSR09/doc/USPolicy_ITUEC.pdf

[2] ITU, "Universal access and service", Module 4. ICT Regulation Toolkit, Section 1.1.2., 2008 [Online]. Available: http://www.ictregulationtoolkit.org/en/Sections.html

[3] UIT, Informe sobre las soluciones innovadoras en materia de gestión y financiación de las políticas de servicio y acceso universales. Ginebra: UIT-D, Comisión de Estudio 1, 3er Periodo de Estudios (2002–2006), 2006 [Online]. Available: http://www.itu.int/opb/publications.aspx?lang=es&parent=D-STGSG01.07.1&folder=D-STG-SG01.07.1-2006

[4] R. del Villar, "Competition and equity in telecommunications", in *No Growth without Equity? Inequality, Interests and Competition in Mexico*, S. Levy and M. Walton, Eds. Washington: Palgrave MacMillan and The World Bank, 2009, pp. 321–364.

[5] C. Casanueva and R. del Villar, "Infrastructure regulation difficulties: the basic telecommunications industry in Mexico (1990–2000)", in *Critical Infrastructures State of the Art in Research and Application*, Wil A. H. Thissen and P. M. Herder, Eds. Dordrecht: Kluwer, 2003, pp. 179–207.

[6] C. Casanueva, "The opening of Mexico's telecommunications markets", in *Telecommunications Reform in India*, R. Dossani. Connecticut: Quorum Books, 2002, pp. 277–296.

[7] R. G. Noll, "Notes on privatizing infrastructure industries", in *Conf. Telecommun. Reform in India*, Asia/Pacific Research Center, Stanford University, November 9 and 10, 2000 [Mimeo].

[8] "Censo de Población y Vivienda de 1990" [Online]. Available: http://www.inegi.org.mx/est/contenidos/proyectos/ccpv/cpv1990/default.aspx

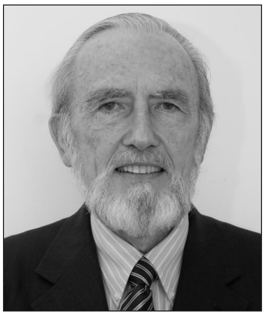
[9] S. Wallsten, S. "Telecommunications privatization in developing countries: the real effects of exclusivity periods", in *Conf. Telecommun. Reform in India*, Asia/Pacific Research Center, Stanford University, November 9 and 10, 2000 [Mimeo].

[10] F. Boehm, "Corrupción y captura en la regulación de los servicios públicos", *Revista de Economía Institucional*, vol. 7, no. 13, pp. 245–262, 2005.

[11] S. Haber, *Crony Capitalism and Economic Growth in Latin America: Theory and Evidence*. Stanford: Hoover Institution, 2002.

Cristina Casanueva-Reguart received her Ph.D. in Development Studies from Stanford University in 1988 and holds an MA from Harvard University and a B.A. in sociology from the Universidad Iberoamericana, Mexico City. She worked in the Federal Government (1992–1998) where she was part of the Federal Competition Commission, Mexico's anti trust agency. Working in the Economic Regulation Department, she focused on the analysis of telecommunications markets, participating in the interministerial group that defined the rules of interconnection. In the following administration she joined the Ministry of Communications, where she was involved in drafting the Federal Law on Telecommunications. In 1998 she returned to an academic position at the Monterrey Institute of Technology, in Mexico City, where she led the creation of an Interdisciplinary Masters Program on the Management of Telecommunications. In 2001, Dr Casanueva joined the Universidad Iberoamericana where she has written numerous articles on telecommunications, regulation and public policy. In 2007–2008, she worked with the Ministry of Communications as a consultant in the field of universal telecommunications service provision.

e-mail: ccasanueva@stanfordalumni.org
Universidad Iberoamericana
Prolongación P. Reforma 880
01210 Mexico D.F., Mexico



Antonio Pita S. obtained his Ph.D. and M.Sc. in solid and fluid mechanics from the University of Iowa in 1967 and 1962, and his B.S.E. from the Johns Hopkins University in 1960. He spent 15 years in teaching and research at the Tecnológico de Monterrey's campus in Monterrey, Mexico and 25 years in R&D at Vitro, a Multinational Mexican-based glass company, including five years as Director of the Corporate R&D Center, 10 years as Technology Director for one of Vitro's Business Units and 10 years as a Senior Researcher. His last assignment was the establishment of Vitro's new R&D centre in Fribourg, Switzerland. Since then he has been Manager of University Relations for the Global Centre for Technology and Innovation that Cemex, a Multinational Mexican-based cement company, has in Brügg, Switzerland and is also Director of the Swiss Office for International Relations for the Tecnológico de Monterrey. He has published 12 articles, 25 Technical reports and holds 3 US and International Patents on Glass and Pre-Reacted Batches of Raw Materials for the Production of Glass Formulas.

e-mail: antonio.pita@itesm.mx
Cemex, Manager of University Relations
for the Global Centre for Technology and Innovation
Römerstrasse 13
2555 Brügg, Switzerland

Paper

Rural Telecommunications Infrastructure Selection Using the Analytic Network Process

Yousef Gasiea, Margaret Emsley, and Ludmil Mikhailov

Abstract—The decisions involved in rural settings are of complex nature, with some aspects compounded by the presence of intangible criteria. Hence, a suitable approach is needed that can produce effective solutions. This paper describes the applicability of a multicriteria decision-making method, specifically the analytic network process (ANP), to model the selection of an appropriate telecommunications infrastructure technology, capable of deploying e-services in rural areas of developing countries. It aims to raise awareness among telecommunication planners about the availability of ANP, and to demonstrate its suitability to enhance the selection process. The proposed model is constructed based on concerned experts' views of relevant selection criteria and potential technology alternatives. Its network structure caters for all possible dependencies and interactions among criteria and alternatives.

Keywords—analytic hierarchy process, analytic network process, multicriteria decision making, rural telecommunications, technology selection.

1. Introduction

Telecommunications technology is evolving rapidly and offers information links between urban and rural areas that can overcome distance barriers and provide e-services to these hardly accessible areas. Recent technological advances in transmission systems like fiber optics, wireless and satellite can now supply services to these locations at affordable prices. However, with different criteria for technology evaluation and various telecommunications infrastructure alternatives available nowadays, the selection process becomes complicated; there is uncertainty and multiple conflicting objectives with sociological, demographical, environmental, political, cultural, economic and technical aspects. This raises the need for some kind of structure or model, based on a suitable multicriteria decision making (MCDM) method.

Some relevant papers cited in literature tackling problems from rural telecommunications field using such methods, with particular focus on the application of analytic hierarchy process (AHP) to rural telecommunications include: Nazem *et al.* [1] use the AHP, to develop a two-phased decision support system to aid the design of rural area telecommunication networks and in [2] examines ways of building an effective rural telecommunications network to facilitate rural development in an information-intensive society. Lee and Kim [3] present a methodology using analytic network process and zero-one goal programming (ZOGP) for information systems projects selection problems that have multiple criteria and interdependence property. In another paper, Nazem *et al.* [4] develops a specific multicriteria decision support mathematical programming model for dealing with the definition of a "hub structure" that is the selection of a number of "nucleus cities" in the context of a rural network planning process. Chemane *et al*. [5] use DecidelT tool based on MCDM to improve the quality of decisions in selecting internet access technologies. Sasidhar and Min [6] use AHP to select the optimal access technology for a rural community under a multiple number of criteria such as cost quality and speed. Nepal [7] applies AHP to the evaluation of rural telecommunications infrastructure. Finally, Andrew *et al*. [8] present a model regarding the applicability of using the AHP for enhancing the selection of communication technologies for rural areas.

While significant decision models are being presented in these papers, but, very few studies have considered all criteria relevant to rural telecommunications, and most of them obviously apply no factor interactions. For example, if a model's emphasis is mainly technical, then the economic, social, regulatory and environmental criteria are probably not adequately addressed. Basically, the AHP is a suitable method when optimization is not pursued, resources are not restricted, and interdependencies between factors do not exist [9]. However, such models do not consider important issues such as interaction among and between decision making levels/clusters as well as dependency among qualitative factors. These are important issues in rural telecommunications decision problems which cannot not be structured hierarchically because they involve many interactions and dependencies requiring a MCDM method to holistically deal with qualitative and quantitative data, with different conflicting objectives, to arrive at a consensus decision in relation to the choice of a suitable rural telecommunication technology.

To the best knowledge of the authors, applications of the analytic network process (ANP) to the selection of rural telecommunications infrastructure technologies have not been cited in the published literature. This paper therefore attempts to fill this gap in the literature to particularly allow for the explicit consideration of dependencies and interactions in the decision making process and still maintains the acknowledged advantages of the AHP method. The ANP is chosen in this paper because of its several advantages over the AHP and other MCDM methods, such as its holistic approach, in which all the factors and criteria involved are laid out in advance in a network system that allows for dependency. Its power lies in its use of special ratio scales to capture interactions for making accurate predictions and reach better decisions [10]. Moreover, its suitability in offering solutions in a complex multicriteria decision environment, together with the availability of software supporting its functions, further acknowledge its applicability to tackle such a problem. It has also proved to be successful in utilizing expert knowledge to tackle several selection problems, e.g., [3] and [11].

The remainder of this paper is organized as follows. Section 2 articulates the selection of rural telecommunications infrastructure problem. The underlying methodology of the proposed approach, the ANP, is briefly introduced in Section 3. The development of the proposed model is explained in Section 4. The pairwise comparisons are described in Section 5. The results are discussed in Section 6 and the paper ends with conclusions in Section 7.

2. The Choice of Technology for Rural Telecommunication Infrastructure

There is a need to provide access to the main telecommunications network and expand connectivity to such areas, thus enabling the rollout of the appropriate telecommunication services. However, the choice of appropriate telecommunications infrastructure technology that will provide the required e-services within various constraints is a challenge. Typically, technology selection is based on a mixture of different criteria, one of which is the remoteness of a village. If the village is within 35 km of the nearest local exchange, telecommunication services can be provided to that village using a one-hop last-mile link. However, if the village is further away, at least two transmission hops must be established [12]. Hence, two types of telecommunications infrastructure technologies are needed to provide rural telecommunication services, namely backbone network (core) and access network (last mile).

The backbone network provides the long-haul signal transmission from the country's main telecommunication centre to the remote access network, i.e., trunking services [12]. This network may be wireless or wireline, including analogue and digital transmission technology over fiber optic, wireless or satellite transmission media [5]. The access network provides the connectivity between the end-user and the backbone network and may be based on wireless or wireline technologies, e.g., copper wires or wireless, connected to network nodes at the edge of the backbone network. Technologies in both networks can be circuit-switched or packet-switched. Any decision made for each of these two segments must take into account the characteristics of rural settlements.

The primary focus of this paper is mainly on the backbone network by attempting to provide a structure of the decision problem and proposing a technology selection model of such an infrastructure. The telecommunications backbone is, in general, a key problem for rural information infrastructure, as low population density is linked to high cost of service for any communications technology, especially for wireline services. It poses the greatest challenge to bringing affordable telecommunication services to rural residents. However, once it is in place and running, it will be possible to connect other nearby rural villages with a wide range of telecommunication technologies and needed services. The infrastructure technology selection process, especially in the case of rural telecommunications in developing countries, is a multi-faceted, multi-criteria decision making problem, requiring consideration of some wide-ranging qualitative factors related to socio-economic and political issues. These are hard to quantify and will have great impact on the selection process, in respect of the social, environmental, regulatory and demographical concerns, etc.

Furthermore, in order to incorporate other tangible factors, in the absence of past statistical data to analyze, such as technical and economic related factors, etc., it is necessary to use a suitable multicriteria method for analysis and synthesis by a group of experts rather than an individual. A telecommunication operator usually receives several technology solutions from external vendors. The challenge of matching the parameters of an engineering problem to the available solutions becomes a challenge to the telecommunications engineer in this particular selection phase [8]. A typical conceptual rural telecommunications infrastructure selection model is illustrated in Fig. 1.

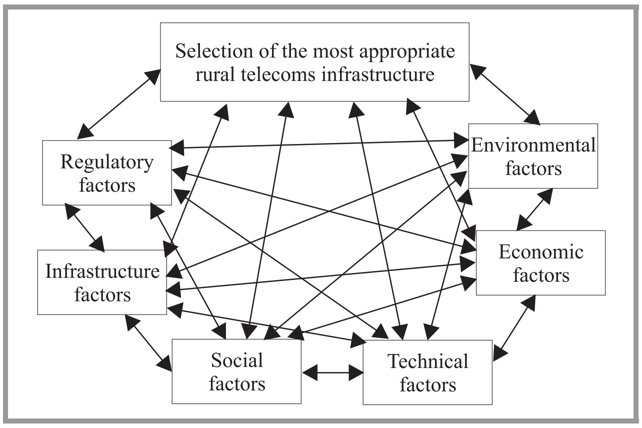


Fig. 1. A conceptual model for the selection of rural telecommunication infrastructure (revised and adapted from [13]).

The obvious significant implication of this conceptual model is that the technical factors are only one subset among others when selecting rural telecommunication technologies, albeit a necessary part. The other factors, such as the sociological, environmental, economic, regulatory and the infrastructure-related are regarded as essential factors that also need to be considered. This can be envisaged as a holistic approach in which the outcome of the selection process is not only dependent on the technical factors, but arises out of the interactions among the various factors. An ANP-based decision model is therefore proposed as a suitable methodology because "*decisions obtained from a network can be significantly different from those obtained from a more complex hierarchy*" [14]. It is constructed to include an in-depth and comprehensive examination of all pertinent factors and will be dependent on the perceptual weightings, provided by telecommunications experts.

3. The Analytic Network Process

The ANP is a multi-attribute decision making approach developed by Thomas L. Saaty and was originally called the supermatrix technique [15]. It is a generalization of the AHP decision methodology where hierarchies are replaced by networks, allowing the capturing of the outcome of dependence and feedback within and among the clusters of elements. Its network structure differs from a hierarchy as illustrated in Fig. 2 [10]. The hierarchy has a goal, levels of elements and connections between the elements.

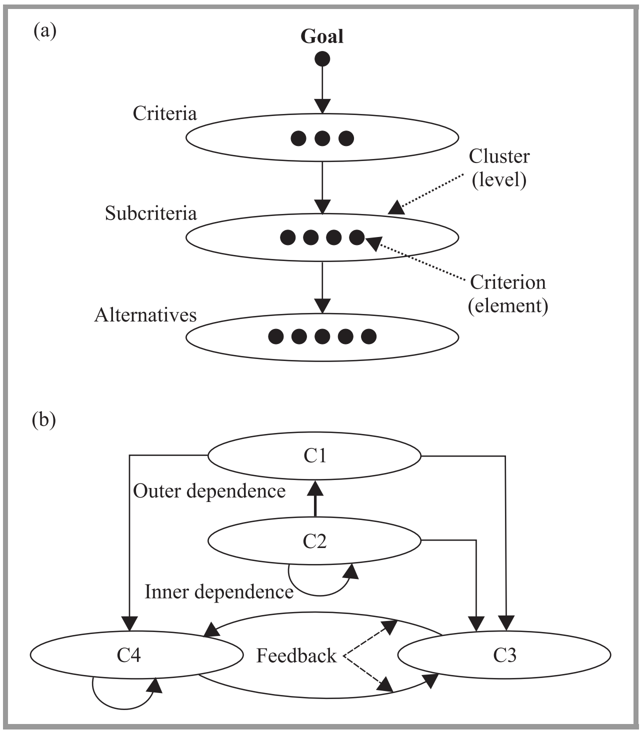


Fig. 2. Examples of a hierarchy (a) and a network (b).

It has no inner dependence and no feedback from lower to higher levels. Unlike the hierarchy, the network structure has no levels but clusters of elements where every element can depend on any other element. The influence is transmitted from one cluster to another (outer dependence) and back, either directly from the second cluster, or, by transiting through intermediate clusters along a path which sometimes can return to the original cluster forming a cycle [10]. The existence of feedback indicates there is mutual outer dependence of criteria in two different clusters, which prevents the problem from being modeled hierarchically due to the difficulty in deciding which cluster is higher/lower than the other. Also, because of inner dependence, the relationships between same level criteria are not represented hierarchically.

The specific ANP model is based on the reasoning, knowledge and experience of experts in the field and relies on the process of eliciting managerial inputs, allowing for a structured communication among decision makers, so that it can act as a qualitative tool for strategic decision-making problems. "*It is a relatively new methodology that is still not well-known to the operations research community and practitioners*" [9]. With its capability to deal with dependence and feedback, it is the most general framework for a detailed analysis of societal, governmental and corporate decisions that is available today to the decision-maker [15]. Therefore, in recent years, there has been an increased use of the ANP in a variety of decision making problems and numerous applications have been published in literature [16].

The ANP is a coupling of two parts. The first part consists of a control network of criteria that controls the interactions in which the criteria should be identified, organized and prioritized in the framework of a control network. The second part is to derive a network of influences among the factors and clusters, i.e., the influence of elements in the feedback system with respect to each of these criteria. Paired comparison judgments of homogeneous elements are performed and synthesized to obtain the priorities of these criteria. The ANP then joins all possible outcomes together in its structures and both judgement and logic are used to estimate the relative influence from which the overall answer is to be derived [15]. The *SuperDecisions* software can be used to perform matrices computation and solve AHP/ANP problems [17].

4. The Development of the Decision Model

In this section, we introduce an ANP model and its development to show how the ANP can be used in the rural telecommunications environment. As each telecommunication infrastructure provider will have its own set of criteria. The attempt here is to present a generalized model based on factors and alternatives identified from the published literature, best practices and telecommunications experts that could then be adapted or extended to support a particular context or a situation of a developing country.

4.1. Setting Selection Criteria

To adapt the ANP methodology for such a technology selection process, it is the foremost activity of the researcher to examine the relevant issues involved. Hence, the first task is the definition of the criteria that will be used for the selection of the appropriate technology for rural connectivity. The activities used to consolidate the final list of the selection criteria, were:

– an intensive literature survey of past studies on similar problems, including: [5], [6], [8], [12] and [18]; the outcome of this activity was the consolidation of an initial list of criteria that comprises the most important factors for the problem at hand;

– interactions with telecommunication experts both from industry and academia from all over the world, who were contacted through e-mail to provide feedback on the initial list of criteria; from the aforementioned activities, a consolidated list of 31 selection criteria, deemed to affect the planners' decision in the choice of rural telecommunications backbone infrastructure, can be observed in Table 1.

Table 1
Ordering and clustering of criteria according to relevance and importance

Cluster	Criteria	Mean
(A) Technical	(A1) Reliability	4.00
	(A2) Ease of maintenance	3.94
	(A3) Remote network management	3.88
	(A4) Compatibility	3.81
	(A5) Ease of installation	3.72
	(A6) Scalability	3.54
	(A7) Bandwidth	3.53
	(A8) Flexibility	3.52
	(A9) Latency	3.30
(B) Infrastructure	(B1) Coverage range	3.80
	(B2) Security of physical infrastructure	3.73
	(B3) Proposed usage	3.40
	(B4) Availability of skilled technicians	3.34
	(B5) Access to existing telecoms infrastructure	3.32
	(B6) Remoteness of area	3.26
	(B7) Rollout time	3.11
	(B8) Parallel infrastructure	2.97
(C) Economic	(C1) Operating cost	4.13
	(C2) Funding sources	4.11
	(C3) Capital cost	3.98
	(C4) Return on investment	3.63
	(C5) Economic development of area	3.32
(D) Social	(D1) Demand	3.77
	(D2) Affordability	3.73
	(D3) Population density	3.48
	(D4) Community of interest	3.42
(E) Regulatory	(E1) Spectrum availability	3.74
	(E2) Licensing constraints	3.52
	(E3) Rights of way	3.30
(F) Environmental	(F1) Terrain topography	3.24
	(F2) Climatic conditions	3.00

4.2. The Online Survey

In order to rank the selection criteria according to their relative importance, an online questionnaire was designed. Telecommunication experts were asked to rate the importance of each factor using a five-point Likert-type scale, ranging from: *not important = 1*, *moderately important = 2*, *strongly important = 3*, *very strongly important = 4* and *extremely important = 5*. A pilot survey was conducted before posting the questionnaire online and subsequently the questionnaire was slightly modified.

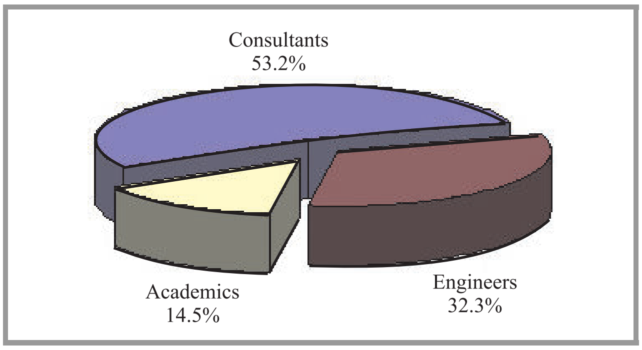


Fig. 3. Categorization of respondents by their professional background.

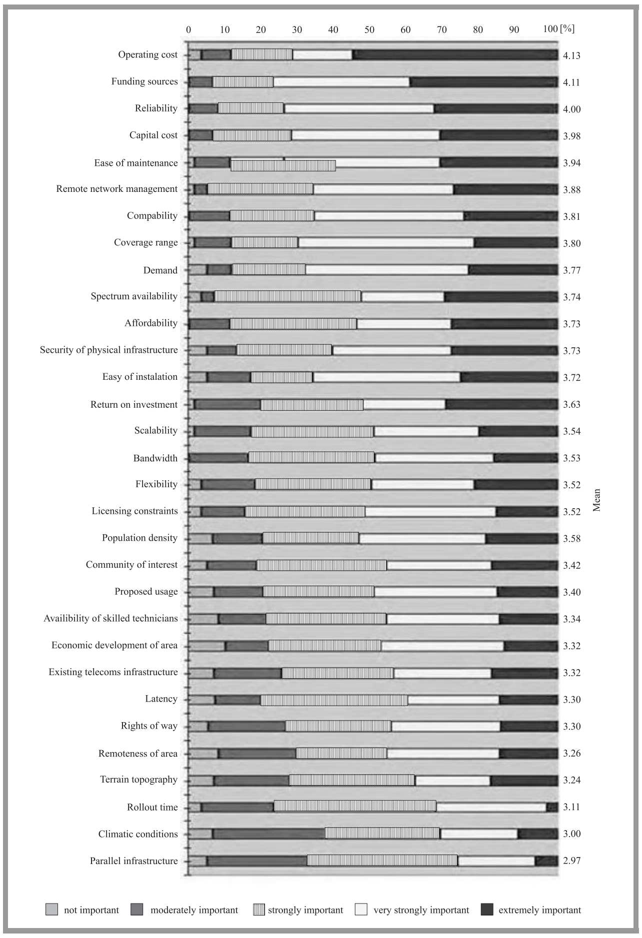


Fig. 4. A graphical representation of the online survey results in percentages.

Table 2
Comparisons of some features of potential alternatives

Alternative	Advantages	Disadvantages
(G1) Fiber optic cable	High speed More reliability High flexibility	High cost Long rollout time Most difficult to deploy
(G2) Power line communication	Simplicity Low cost Use of power lines	Less reliability Data signal disruption Noise and interference
(G3) Microwave link	High speed Low cost equipment Fast deployment	Low reach and line of sight Licensing constraints Less bandwidth and flexibility
(G4) Satellite communication	Wide coverage Ease of deployment Overcomes topography	High latency High cost Limited bandwidth

The obtained responses effectively reached 62 responses, which is considered adequate because the purpose of the survey was mainly to obtain a range of diversified expert opinions with respect to each particular selection factor.

The respondents' profiles showed that all of them are generally involved in telecommunications field, where some of them are particularly dealing with rural telecommunications projects. They can be categorized by their professional backgrounds into three categories as shown in Fig. 3: of the 62 respondents 20 (32.3%) of them work as telecommunication engineers, 33 (53.2%) as consultants and 9 (14.5%) as academics. This mix up of the respondents' expertise confirms their familiarity with the selection factors and also indicates that they were very well placed to provide useful data for such a survey.

The results were then analyzed using SPSS (SPSS Inc., 2006), and univariate descriptive statistics were generated, including the relative importance index for each factor. Figure 4 summarizes the obtained results and shows that all proposed criteria are mostly within the *strongly important* and *very strongly important* categories, the only exception being the results of the "operating cost" criterion which is inclined more towards the *extremely important* grade.

4.3. Grouping of Criteria into Clusters

The mean rating values were used to group the criteria into six clusters coded A through F according to relevance, in this order: (A) Technical, (B) Infrastructure, (C) Economic, (D) Social, (E) Regulatory, and (F) Environmental. Each cluster only includes criteria that are comparable or do not differ by orders of magnitude [10]. Table 1 shows the coding and the ordering of criteria for all clusters.

4.4. Alternatives Identification

The activities abovementioned in Subsection 4.1 were repeated in order to identify potential technology alternatives. The published literature, e.g., [19], identified four technological solutions to provide rural backbone infrastructure to promote e-services in rural areas of developing countries that include two wireline technologies: fiber optic cable and power line communication, and two wireless technologies: fixed wireless and satellite, which were initially highlighted as candidate decision alternatives for this research. After consultation with telecommunication experts, the alternatives finally selected for this research are (G1) fiber optic cable, (G2) power line communication, (G3) microwave link and (G4) satellite communication. Table 2 briefly summarizes some characteristics of alternatives.

4.5. Assessing Dependencies

After structuring the decision problem, the next step is to examine the dominance of influence among criteria. In order to fulfil this task, a new survey questionnaire was distributed to experts who had an overview of the research, were interested and actually involved in the field of rural telecommunications, who were asked to identify the dependencies among criteria. Seven completed questionnaires were collected. The majority rule was then used to aggregate the responses into a single matrix, which was developed using a zero-one matrix of criteria against criteria using a binary value of 1 to signify dependence of one criterion on another, and zero otherwise [20]. A majority condition of 4 out of 7 (4/7) experts' consensus (i.e., 57%) was considered as a minimum requirement for any entry that indicates the existence of a direct relationship between any pair of criteria.

Table 3 shows all possible connections, where the entries can take the following values:

0 indicates no relationship exists based on 7 experts' consensus;

0 indicates the entries have obtained < 4 experts' consensus;

1 indicates the entries have obtained ≥ 4 experts' consensus.

Table 3
The aggregated dependency matrix showing connections among all elements

Clusters	Criteria \ with respect to	A1	A2	A3	A4	A5	A6	A7	A8	A9	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	D1	D2	D3	D4	E1	E2	E3	F1	F2	G1	G2	G3	G4
A	A1	0	1	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
	A2	0	0	1	1	1	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
	A3	0	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	A4	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
	A5	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
	A6	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	1	1	0	1	1	1	1	1	1
	A7	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	1	1	1	1
	A8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1	1	1	1
	A9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	1
B	B1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1
	B2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1	1	1
	B3	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1
	B4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	B5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	1	1	1
	B6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	1
	B7	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1
	B8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1
C	C1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1
	C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1
	C3	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	1	1	1
	C4	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	1	0	1	0	1	1	1	1	1	0	0	0	1	0	1	1	1	1
	C5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1
D	D1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1
	D2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1
	D3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
	D4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
E	E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1	1
	E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1
	E3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1
F	F1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	F2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
G	G1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
	G2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
	G3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
	G4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0

As a result, the entries represented by 1 indicate the existence of a direct relationship from criterion i to criterion j based on the consensus of at least 4 experts, i.e., if criterion i depends on criterion j, the entry a_{ij} will take 1. The criteria in the rows are evaluated with respect to the criteria in the columns, i.e., the 1 in the columns will determine which criteria in the rows are to be pairwise compared with respect to that column. Subsequently, a pairwise comparison matrix will be constructed only for the dependent criteria. Using the *Design module* of the *SuperDecisions* software [17], the network model was constructed according to Table 3, the connections between clusters are illustrated in Fig. 5. A cluster is connected to another cluster when at least one element in it is connected to at least two elements in another cluster. It should be noted that two-way arrows connecting the clusters represent interdependencies among elements, where an arrow direction signify dependence and starts from an element to another that may influence it [17].

Figure 5 contains the entire inner dependence – the parent element and the elements to be compared are in the same cluster so that the cluster is linked to itself and a loop link appears – among elements within each cluster except in the environmental and alternative clusters. It indicates that the connections between the elements are in the same cluster. For example, column A8 means A2, A4, A5, A6 are interrelated with respect to A8.

The proposed model also contains outer dependence which is the relationship between elements in one cluster with others in other clusters [15]. For example, in Table 3, when considering A8, the elements G1, G2 G3 and G4 in the (G) *Alternative cluster* are interconnected and pairwise compared with respect to A8 in the (A) *Technical cluster*. The exception is the regulatory and environmental

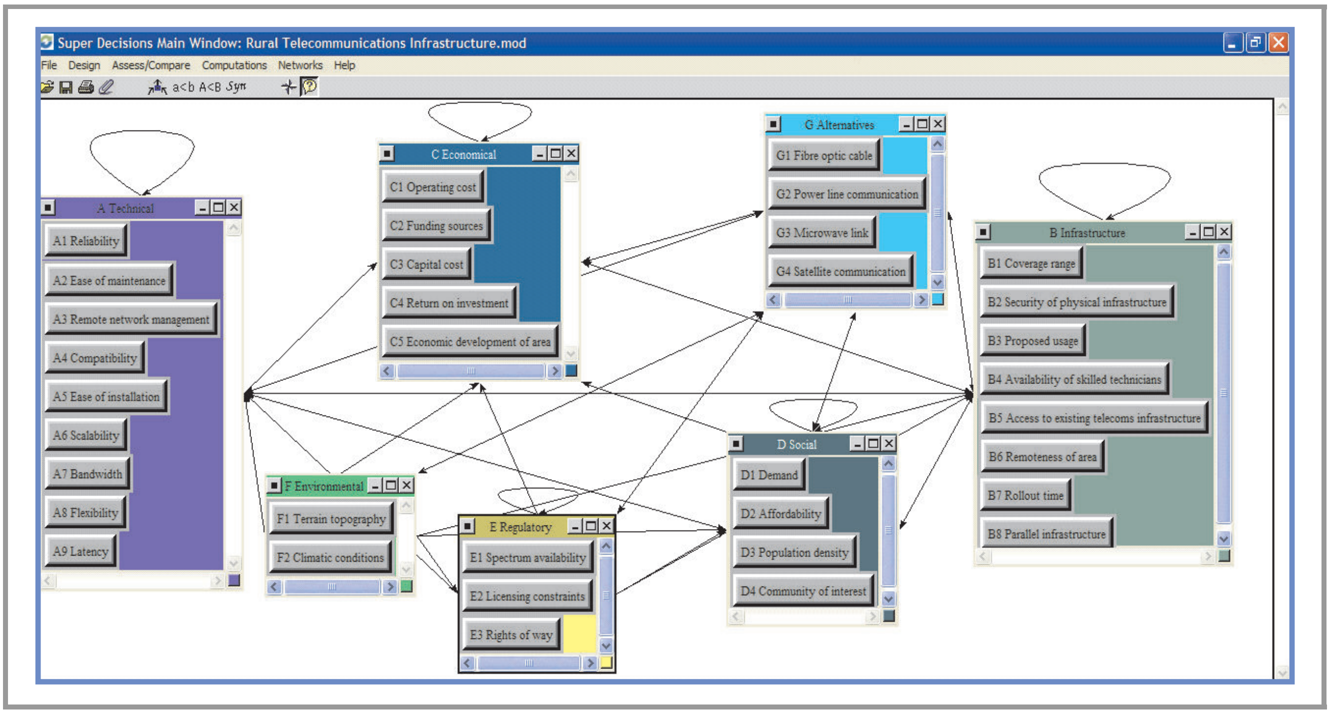


Fig. 5. The ANP network model with connections among elements/clusters.

clusters, i.e., none of the elements in both clusters depend on elements from other clusters with respect to a common attribute within a cluster. Also, the technical and social clusters have no outer dependence on the economic cluster.

Feedback links in which one compares the alternatives with respect to criteria, as in a hierarchy, and also compares the dominance of one criterion versus another for each alternative exist in this structure. Table 3 illustrates that there is mutual outer dependence of criteria in two different clusters as can be seen between the alternative cluster and all other clusters; technical and social clusters and infrastructure and economic clusters. For example, G1 is the parent element and all elements in other clusters except in cluster G are its children elements, which indicates that criteria may be compared with respect to an alternative. This is the strength of the ANP approach because dependence and feedback are incorporated in real life problems, in which a decision process not only compares alternatives with respect to criteria but also vice versa. For instance, in addition to separately comparing G1, G2, G3 and G4 with respect to A1 and A7, A1 and A7 must also be compared with respect to G1. A pairwise question to be asked is: what is a more dominant characteristic of fiber optic cable technology, its reliability or its bandwidth? However, since feedback involves cycles, and cycling can be an infinite process, the operations needed to derive the priorities become more demanding than with hierarchies [20].

Based on the above analysis, it is obvious that the developed inner and outer dependence and feedback among the network structure shown in Table 3 excludes the hierarchy form and calls for the network form to model the selection of rural telecommunications infrastructure technology.

5. Pairwise Comparisons

After constructing the ANP network, the next phase is the measurement and data collection stage which involves compiling a list of experts to provide judgements for pairwise comparisons. Both the AHP/ANP derive ratio scale priorities by making paired comparison of elements on common elements. The subjective judgements are to be entered and assigned a numerical value based on the nine-point scale suggested by Saaty [21] to obtain the corresponding pairwise judgment matrices. A score of 1 indicates the equality between the two elements whereas score 9 represents the dominance of the row element in the matrix over the column element. A reciprocal value is automatically assigned in the opposite position in the matrix, i.e., $a_{ij} = 1/a_{ij}$.

In this model, pairwise comparisons are identified according to the connections developed in Table 3 and then relevant pairwise comparison matrices are created accordingly. The columns in the table present the parent elements, while the rows present the children elements in the structure. For example, G1 is a parent element and A1 through F2 are its children elements. The elements that are to be pairwise compared are always all in the same cluster. They are compared with respect to their parent element, the element from which they are connected.

There are a number of comparison matrices for every parent element, and one comparison matrix for elements in the same cluster originating from the same parent element.

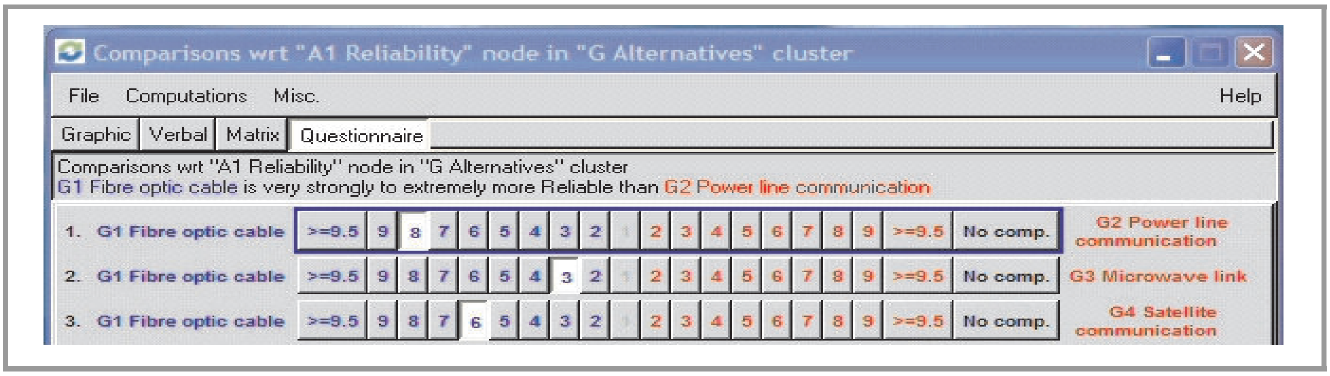


Fig. 6. An example of *SuperDecisions* pairwise comparison process.

For example, there are four comparison matrices for criterion B6, one for each of clusters A, B, C and G. Elements within D cluster cannot be compared with respect to B6 because there should be at least two entries of 1 available within any cluster to perform pairwise comparisons. Therefore, A1, A2, and A5 through A9 are pairwise compared with respect to B6; B1 through B5, B7 and B8 are pairwise compared with respect to B6; C1 and C5 are pairwise compared with respect to B6; and G1 through G4 are pairwise compared with respect to B6. This results in local priorities of the children elements with respect to the parent element. It is only necessary to make $n(n-1)/2$ comparisons to establish the full set of pairwise judgements, where n denotes the number of elements (nodes). For example, six pairwise comparison questions are required for A1 because $n = 4$ for the alternatives outer dependence on A1, while for A8, twelve pairwise comparison questions are needed because $n = 4$ for the inner dependence within the technical cluster, and also $n = 4$ for the alternatives outer dependence on A8.

In this developed structure; there are a total of 92 judgement matrices which include 674 pairwise comparison questions for both inner and outer dependences developed within the network. It is obvious that the task of asking such a large number of questions would be very enormous and would require intensive efforts and extended time. Hence, in order to establish a more rational approach to collect pairwise comparison judgements from qualified telecommunication experts, and also to economize efforts, it was decided to design and use several online questionnaires to gather data from experts. The questionnaires included all required pairwise questions to assess expert judgments in relation to the relative influence of affecting elements on the affected ones.

An example of such pairwise question is: "In selecting an appropriate backbone infrastructure technology in rural areas of developing countries, which influences fiber optic cable technology more, ease of installation or ease of maintenance? Conversely, given the ease of installation, which of these technologies are more dominant, fiber optic cable or satellite?"

Since the clusters in this network are not equally important, their weights in the cluster matrix are obtained by pairwise comparisons. Each cluster is taken in turn as a parent cluster, and the other clusters connected to it are pairwise compared for importance with respect to their influence on it [17]. It should be noted that the pairwise comparisons to assess the influence of some cluster on all other clusters is actually what distinguishes the ANP from the AHP.

For example, one of the cluster comparison questions addressed to the experts is: "Which influences the selection of rural telecommunications backbone infrastructure more, economic or technical issues?" The obtained cluster weights are used in a later stage to weight all the elements in the unweighted supermatrix. The individual expert pairwise comparisons are aggregated into a representative group judgment, by applying geometric means.

A score corresponding to the group judgment regarding this question is then clicked to highlight the technology providing more reliability relative to the technology providing less reliability. While, a score of 1 indicates the equality between the two technologies, the blue scores represent the dominance of the row element in the matrix (e.g., G1) over the column element (e.g., G2) and the red scores are vice versa. A reciprocal value is automatically assigned in the opposite position in the matrix.

An example of the comparison process used in *SuperDecisions* is shown in Fig. 6. It presents the pairwise comparisons between alternatives G1 and G2, regarding the reliability factor. The question being asked is "With respect to reliability, which technology is more reliable: fiber optic technology or power line communication?" The group judgment was that G1 is between *very strongly* and *extremely* more reliable than G2, therefore the comparison value of 8 is entered.

The comparison between all other alternatives regarding different criteria is done in the same way.

The next stage of the process includes the computations of the relative importance of the elements. For each comparison matrix a local priority vector (also referred as an eigenvector) is computed, by applying the eigenvector approach [14], provided that the inconsistency ratio (IR) of this matrix is less than 0.1.

Table 4
Comparison matrix of alternatives with respect to reliability, $IR = 0.0958$

Alternative	G1	G2	G3	G4	Eigenvector
(G1) Fiber optic cable	1.000	8.240	3.350	5.730	**0.580**
(G2) Power line communication	0.121	1.000	0.178	0.217	**0.044**
(G3) Microwave link	0.299	5.630	1.000	3.830	**0.262**
(G4) Satellite communication	0.175	4.610	0.261	1.000	**0.114**

Table 5
Ordering and clustering of criteria according to relevance and importance

Cluster	Criteria	Priorities [%]	
		Normalized	Limiting
(A) Technical	(A1) Reliability	18.73	2.42
	(A2) Ease of maintenance	21.70	2.81
	(A3) Remote network management	17.05	2.21
	(A4) Compatibility	2.62	0.33
	(A5) Ease of installation	2.80	0.36
	(A6) Scalability	11.98	1.55
	(A7) Bandwidth	13.42	1.74
	(A8) Flexibility	3.50	0.45
	(A9) Latency	8.21	1.06
(B) Infrastructure	(B1) Coverage range	58.32	7.96
	(B2) Security of physical infrastructure	5.04	0.69
	(B3) Proposed usage	12.07	1.65
	(B4) Availability of skilled technicians	3.24	0.44
	(B5) Access to existing telecoms infrastructure	4.74	0.65
	(B6) Remoteness of area	3.22	0.44
	(B7) Rollout time	5.84	0.80
	(B8) Parallel infrastructure	7.53	1.03
(C) Economic	(C1) Operating cost	16.39	7.53
	(C2) Funding sources	34.70	15.94
	(C3) Capital cost	7.84	3.60
	(C4) Return on investment	37.32	17.15
	(C5) Economic development of area	3.75	1.72
(D) Social	(D1) Demand	64.96	1.74
	(D2) Affordability	20.49	0.55
	(D3) Population density	9.91	0.27
	(D4) Community of interest	4.65	0.12
(E) Regulatory	(E1) Spectrum availability	56.99	0.67
	(E2) Licensing constraints	27.45	0.32
	(E3) Rights of way	15.56	0.18
(F) Environmental	(F1) Terrain topography	56.81	0.71
	(F2) Climatic conditions	43.19	0.54

The *SuperDecisions* can also deal with the issue of improving consistency of the matrices, by identifying the most inconsistent judgments. The matrix consistency can then be improved by providing more consistent judgements by the decision makers so that $IR \leq 0.12$. For further explanation of inconsistency and how to calculate it, one can refer to [22].

An example of an aggregated comparison matrix within the alternative cluster (G) with respect to reliability (A1), and the corresponding values of the eigenvector, is shown in Table 4.

From this table we can see that in terms of reliability, the fiber optic cable technology has the highest priority (0.580) followed by microwave links and satellite with (0.262) and (0.1244), respectively. The less reliable technology is the power line communication (0.044). Since IR is less than 0.12, this matrix is considered of acceptable consistency.

The eigenvector derived in this way is then entered as a part of some column of a supermatrix. It represents the impact of a given set of elements in a component on another element in the system, where a component in a supermatrix is the block, defined by a cluster name at the left and a cluster name at the top. If an element has no influence on another element, its influence priority is assigned zero [20]. The formation of a supermatrix in the ANP allows for the resolution of the effects of the interdependence that exists between the elements of the system.

The *SuperDecisions* performs necessary matrix operations for structuring of the three supermatrices, associated with this model, as shown in the Appendix. Table A1 illustrates the unweighted supermatrix that contains the local priorities derived from pairwise comparisons throughout the network; they can be read directly from this matrix. The weighted supermatrix shown in Table A2; is obtained by multiplying all the elements in a component of the unweighted supermatrix by the corresponding cluster weight, i.e., each block of column eigenvectors belonging to a component is weighted by the priority of influence of that component. This makes the entire columns sum to unity exactly, i.e., the weighted supermatrix is said to be "column stochastic". Finally, the limit supermatrix is obtained by raising the weighted supermatrix to the power k, where k is an arbitrarily large number, to allow for convergence of the interdependent relationships.

The final values of priorities of all the elements are obtained by normalising each block, so that the columns of the limit supermatrix become identical. The values of the priorities of all elements can be read from any column [15] as can be seen in Table A3.

The *SuperDecisions* has also been used to produce the priorities shown in Table 5. It contains the relative importance of all criteria considered in the model. For example, under the limiting priorities' column, one can observe that the most important factors among all are the *Return on investment* criterion with a priority of 17.15% followed by the *Funding sources* criterion with 15.94%. According to the *Normalized* priorities column, the most important criterion is the *Demand* with a priority of 64.96%, followed by the *Coverage range* with 58.32%. Among the technical criteria; the *Ease of maintenance*, *Reliability* and *Remote network management* criteria have the highest priorities of 21.71%, 18.73% and 17.05%, respectively. The *Spectrum availability* and *Terrain topography* factors are regarded as the most important within regulatory and environmental clusters, with priorities of 56.99% and 56.81%, respectively.

The relative importance of all other criteria considered in the model can be seen in Table 5.

6. Conclusions

This research paper reports on the applicability of using a MCDM method to enhance the selection process of an essential rural infrastructure technology. An ANP model incorporates both qualitative and quantitative approaches to a decision problem. The qualitative part includes:

- identification of the decision problem;
- ensuring the suitability of ANP to solve the problem;
- decomposing the unstructured problem to a set of manageable and measurable levels;
- compiling a list of experts to provide judgements for making the decision.

The quantitative part includes:

- designing a questionnaire to collect input data through pairwise comparison;
- estimating the relative importance between any two elements in each matrix and calculating the relevant eigenvectors;
- measuring the inconsistency of each matrix by employing the consistency ratio;
- eventually constructing the supermatrix using the eigenvectors of the individual matrices.

Based on the performed analysis, it is shown that the problem has inner and outer dependences and feedback among the elements, which excludes the hierarchy form (AHP) and requires a network form to model the selection process. The paper illustrates the use of the ANP method, but no real life conclusions should be drawn from it, as each telecommunication infrastructure provider will have its own set of criteria. The attempt here is to present a generic model based on factors and alternatives identified from the published literature, best practices and telecommunications experts that could then be adapted or extended to support a particular context or a situation of a developing country. Planners may therefore augment this model with their own company-specific factors that might change the priorities.

The obtained results reflect the preferences of experts who made the judgments, therefore, they cannot be considered as an objective assessment of the relative suitability of the four technologies as backbone infrastructure in rural areas. Final alternatives scores should, therefore, be thought of as an input to the decision-making process rather than its end. This process would be refined with experience, optimising the accuracy and time taken to reach proper decisions regarding the choice of telecommunication infrastructure in rural surroundings.

Appendix

Table A1
The unweighted supermatrix

		A									B								C					D				E			F		G			
		A1	A2	A3	A4	A5	A6	A7	A8	A9	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	D1	D2	D3	D4	E1	E2	E3	F1	F2	G1	G2	G3	G4
A	A1	0.000	1.000	0.000	0.000	0.000	0.000	0.631	0.000	0.000	0.000	0.815	0.000	0.132	0.000	0.125	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.113	0.463	0.139	0.046	0.151	0.019
	A2	0.000	0.000	1.000	0.569	0.543	0.000	0.000	0.416	0.000	0.000	0.101	0.000	0.719	0.000	0.157	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.768	0.393	0.289	0.151	0.042	0.098	0.029
	A3	0.000	0.000	0.000	0.431	0.000	0.000	0.291	0.000	0.000	0.721	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.102	0.046	0.084	0.050
	A4	0.000	0.000	0.000	0.000	0.457	0.000	0.000	0.091	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.039	0.071	0.236	0.024	0.050
	A5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.080	0.000	0.000	0.083	0.000	0.150	0.000	0.527	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.232	0.393	0.169	0.052	0.181	0.028	0.044
	A6	0.000	0.000	0.000	0.000	0.000	0.000	0.078	0.413	0.000	0.000	0.000	0.000	0.000	0.000	0.039	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.543	0.000	0.123	0.000	0.146	0.208	0.000	0.035	0.040	0.063	0.249	0.025	0.118
	A7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.279	0.000	1.000	0.000	0.000	0.029	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.457	0.000	0.777	0.000	0.774	0.424	0.000	0.000	0.000	0.380	0.032	0.131	0.112
	A8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.000	0.203	0.000	0.032	0.000	0.025	0.149	0.055	0.245
	A9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.080	0.166	0.000	0.000	0.000	0.017	0.018	0.404	0.332
B	B1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.471	0.000	0.168	0.000	0.000	0.000	1.000	0.396	0.155	0.000	0.572	0.708	0.510	0.644	0.000	0.587	0.000	0.231	0.083	0.402	0.030
	B2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.148	0.000	0.000	0.000	0.000	0.000	0.000	0.062	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.000	0.082	0.059	0.133	0.230
	B3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.163	0.000	0.030	0.000	0.000	0.000	0.000	0.000	0.000	0.056	0.000	1.000	0.265	0.137	0.067	0.298	0.000	0.036	0.000	0.092	0.173	0.046	0.074
	B4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.030	0.028	0.106	0.154
	B5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.173	0.000	0.832	0.000	0.000	0.000	0.000	0.123	0.000	0.000	0.000	0.000	0.000	0.000	0.763	0.000	0.000	0.023	0.103	0.156	0.034
	B6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.090	0.155	0.000	0.000	0.000	0.117	0.226	0.213	0.105	0.094	0.022
	B7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.837	0.000	0.065	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.845	0.000	0.073	0.000	0.424	0.057	0.237	0.225	0.774	0.304	0.018	0.038	0.052
	B8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.075	0.000	0.000	0.000	0.000	0.000	0.000	0.363	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.430	0.026	0.404
C	C1	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.838	0.000	0.245	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.310	0.000	1.000	1.000	1.000	0.427	0.873	0.563	0.046	0.348	0.077
	C2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.744	1.000	0.431	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.155	0.154	0.404	0.354
	C3	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.265	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.127	0.000	0.058	0.000	0.000	0.000	0.000	0.000	0.000	0.113	0.507	0.043	0.064
	C4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.735	0.000	1.000	0.000	0.000	0.000	0.000	0.755	1.000	0.000	0.256	0.000	0.569	0.873	1.000	0.063	1.000	0.000	0.000	0.000	0.093	0.000	0.127	0.144	0.102	0.284
	C5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.162	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.570	0.000	0.000	0.000	0.000	0.480	0.127	0.042	0.147	0.103	0.221
D	D1	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.551	0.423	0.672	0.211
	D2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	0.000	0.000	0.304	0.412	0.168	0.560
	D3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.071	0.089	0.117	0.078
	D4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.074	0.076	0.043	0.152
E	E1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.500	0.000	0.480	0.065	0.708	0.602
	E2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.424	0.227	0.219	0.330
	E3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	0.095	0.709	0.073	0.067
F	F1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.891	0.192	0.843	0.124
	F2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.109	0.808	0.157	0.876
G	G1	0.580	0.083	0.115	0.149	0.126	0.431	0.615	0.204	0.473	0.058	0.390	0.483	0.471	0.064	0.052	0.483	0.535	0.047	0.104	0.048	0.084	0.309	0.561	0.044	0.064	0.048	0.051	0.290	0.579	0.043	0.533	0.000	0.000	0.000	0.000
	G2	0.044	0.113	0.082	0.056	0.060	0.054	0.064	0.057	0.341	0.081	0.436	0.356	0.380	0.090	0.071	0.385	0.311	0.593	0.082	0.243	0.122	0.508	0.275	0.078	0.105	0.283	0.052	0.520	0.300	0.072	0.315	0.000	0.000	0.000	0.000
	G3	0.262	0.494	0.248	0.572	0.206	0.401	0.257	0.288	0.148	0.261	0.100	0.071	0.053	0.559	0.318	0.087	0.088	0.226	0.487	0.537	0.618	0.117	0.114	0.598	0.234	0.063	0.293	0.126	0.070	0.288	0.093	0.000	0.000	0.000	0.000
	G4	0.114	0.310	0.555	0.224	0.607	0.114	0.064	0.451	0.038	0.599	0.074	0.090	0.096	0.288	0.560	0.045	0.066	0.134	0.327	0.171	0.177	0.065	0.049	0.280	0.597	0.606	0.604	0.063	0.052	0.597	0.059	0.000	0.000	0.000	0.000

Table A2
The weighted supermatrix

		A									B								C					D				E			F		G			
		A1	A2	A3	A4	A5	A6	A7	A8	A9	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	D1	D2	D3	D4	E1	E2	E3	F1	F2	G1	G2	G3	G4
A	A1	0.000	0.403	0.000	0.000	0.000	0.000	0.428	0.000	0.000	0.000	0.160	0.000	0.023	0.000	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.082	0.023	0.008	0.025	0.003
	A2	0.000	0.000	0.869	0.495	0.219	0.000	0.000	0.362	0.000	0.000	0.020	0.000	0.125	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.194	0.066	0.051	0.025	0.007	0.016	0.005
	A3	0.000	0.000	0.000	0.375	0.000	0.000	0.197	0.000	0.000	0.141	0.000	0.000	0.000	0.534	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.007	0.014	0.008
	A4	0.000	0.000	0.000	0.000	0.184	0.000	0.000	0.079	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.007	0.012	0.038	0.004	0.008
	A5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.070	0.000	0.000	0.016	0.000	0.026	0.000	0.084	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.059	0.066	0.030	0.008	0.030	0.005	0.007
	A6	0.000	0.000	0.000	0.000	0.000	0.000	0.053	0.359	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.816	0.000	0.000	0.000	0.000	0.000	0.000	0.162	0.000	0.035	0.000	0.037	0.053	0.000	0.006	0.007	0.010	0.041	0.004	0.019
	A7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.055	0.000	0.178	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.136	0.000	0.219	0.000	0.194	0.107	0.000	0.000	0.000	0.062	0.005	0.021	0.018
	A8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.000	0.000	0.051	0.000	0.005	0.000	0.004	0.024	0.009	0.040
	A9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.042	0.000	0.000	0.000	0.003	0.003	0.066	0.054
B	B1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.346	0.049	0.000	0.023	0.000	0.000	0.000	0.360	0.142	0.028	0.000	0.099	0.170	0.091	0.116	0.000	0.073	0.000	0.046	0.016	0.079	0.006
	B2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.016	0.012	0.026	0.045
	B3	0.152	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.683	0.000	0.000	0.000	0.018	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.000	0.240	0.046	0.033	0.012	0.054	0.000	0.004	0.000	0.018	0.034	0.009	0.015
	B4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.021	0.030
	B5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.113	0.000	0.000	0.000	0.000	0.044	0.000	0.000	0.000	0.000	0.000	0.000	0.137	0.000	0.000	0.005	0.020	0.031	0.007
	B6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.037	0.000	0.000	0.000	0.015	0.030	0.042	0.021	0.018	0.004
	B7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.094	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.154	0.000	0.013	0.000	0.076	0.010	0.043	0.028	0.102	0.060	0.004	0.007	0.010
	B8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.131	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.085	0.005	0.080
C	C1	0.625	0.536	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.674	0.000	0.520	0.000	0.200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.137	0.000	0.460	0.463	0.463	0.225	0.486	0.262	0.022	0.162	0.036
	C2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.661	0.569	0.245	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.072	0.072	0.188	0.165
	C3	0.000	0.000	0.000	0.000	0.536	0.000	0.000	0.000	0.000	0.201	0.760	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.060	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.053	0.237	0.020	0.030
	C4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.559	0.000	0.691	0.000	0.000	0.000	0.000	0.616	0.888	0.000	0.227	0.000	0.324	0.408	0.614	0.028	0.614	0.000	0.000	0.000	0.049	0.000	0.059	0.067	0.048	0.132
	C5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.251	0.000	0.000	0.000	0.000	0.253	0.071	0.019	0.069	0.048	0.103
D	D1	0.152	0.000	0.000	0.000	0.000	0.000	0.220	0.000	0.000	0.000	0.000	0.090	0.000	0.000	0.081	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.079	0.057	0.000	0.000	0.000	0.000	0.000	0.000	0.037	0.029	0.045	0.014
	D2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.079	0.065	0.000	0.000	0.000	0.000	0.021	0.028	0.011	0.038
	D3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.093	0.098	0.005	0.006	0.008	0.005
	D4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.005	0.003	0.010
E	E1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.057	0.000	0.026	0.000	0.024	0.003	0.035	0.030
	E2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.057	0.000	0.000	0.021	0.011	0.011	0.016
	E3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.026	0.000	0.005	0.035	0.004	0.003
F	F1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.050	0.011	0.047	0.007
	F2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.045	0.009	0.049
G	G1	0.041	0.005	0.015	0.019	0.008	0.431	0.063	0.027	0.150	0.003	0.017	0.019	0.018	0.008	0.002	0.089	0.025	0.005	0.104	0.005	0.006	0.022	0.029	0.003	0.003	0.003	0.002	0.014	0.027	0.002	0.020	0.000	0.000	0.000	0.000
	G2	0.003	0.007	0.011	0.007	0.004	0.054	0.007	0.007	0.108	0.004	0.019	0.014	0.015	0.011	0.003	0.071	0.015	0.066	0.082	0.027	0.009	0.036	0.014	0.005	0.005	0.019	0.002	0.024	0.014	0.003	0.012	0.000	0.000	0.000	0.000
	G3	0.019	0.030	0.032	0.075	0.012	0.401	0.026	0.038	0.047	0.012	0.004	0.003	0.002	0.067	0.011	0.016	0.004	0.025	0.487	0.060	0.044	0.008	0.006	0.040	0.011	0.004	0.014	0.006	0.003	0.010	0.003	0.000	0.000	0.000	0.000
	G4	0.008	0.019	0.072	0.029	0.037	0.114	0.007	0.059	0.012	0.026	0.003	0.004	0.004	0.035	0.020	0.008	0.003	0.015	0.327	0.019	0.013	0.005	0.003	0.019	0.029	0.041	0.028	0.003	0.002	0.021	0.002	0.000	0.000	0.000	0.000

Table A3
The limit supermatrix

		A									B								C					D				E			F		G			
		A1	A2	A3	A4	A5	A6	A7	A8	A9	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	D1	D2	D3	D4	E1	E2	E3	F1	F2	G1	G2	G3	G4
A	A1	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242
	A2	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281	0.0281
	A3	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221	0.0221
	A4	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034
	A5	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036
	A6	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155
	A7	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174
	A8	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045	0.0045
	A9	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106	0.0106
B	B1	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796	0.0796
	B2	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069
	B3	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165
	B4	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044
	B5	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065
	B6	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044	0.0044
	B7	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080	0.0080
	B8	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103	0.0103
C	C1	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753	0.0753
	C2	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594	0.1594
	C3	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360	0.0360
	C4	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715	0.1715
	C5	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172
D	D1	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174	0.0174
	D2	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055
	D3	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
	D4	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012
E	E1	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067
	E2	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032
	E3	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018
F	F1	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071	0.0071
	F2	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054	0.0054
G	G1	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322
	G2	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259
	G3	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013	0.1013
	G4	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642	0.0642

References

[1] S. M. Nazem, H. Lee, and Y. Shi, "Designing rural area telecommunication networks via hub cities", *Omega*, vol. 22, pp. 305–314, 1994.

[2] S. M. Nazem, Y. H. Liu, H. Lee, and Y. Shi, "Implementing telecommunications infrastructure: a rural America case", *Telemat. Inform.*, vol. 13, pp. 23–31, 1996.

[3] J. W. Lee and S. H. Kim, "Using analytic network process and goal programming for interdependent information system project selection", *Comput. Oper. Res.*, vol. 27, pp. 367–382, 2000.

[4] S. M. Nazem, Y. Shi, H. Lee, K. Sung Yeol, P. Tae Ho, and S. Myung Ho, "Multicriteria hub decision making for rural area telecommunication networks", *Eur. J. Oper. Res.*, vol. 133, no. 3, pp. 483–495, 2001.

[5] L. Chemane, L. Ekenberg, O. Popov, and T. Cossa, "MCDM model for selecting internet access technologies – a case study in Mozambique", in *IEEE EUROCON 2005*, Belgrade, Serbia & Montenegro, 2005.

[6] M. Sasidhar and K. J. Min, "Decision support models for the selection of internet access technologies in rural communities", *Telemat. Inform.*, vol. 22, no. 3, pp. 201–219, 2005.

[7] T. Nepal, "Evaluation of rural telecommunications infrastructure in South Africa", *Probl. Nonlin. Anal. Eng. Syst.*, vol. 3, no. 24, pp. 138–149, 2005.

[8] T. N. Andrew, P. Rahoo, and T. Nepal, "Enhancing the selection of communication technology for rural telecommunications: an analytic hierarchy process model", *Int. J. Comput. Syst. Sig.*, vol. 6, no. 2, 2005.

[9] J. S. Shang, Y. Tjader, and Y. Ding, "A unified framework for multicriteria evaluation of transportation projects", *IEEE Trans. Eng. Manage.*, vol. 51, no. 3, 2004.

[10] T. L. Saaty, *Theory and Applications of the analytic Network Process: Decision Making with Benefits, Opportunities, Costs, and Risks*. Pittsburgh: RWS Publications, 2005.

[11] L. M. Meade and A. Presley, "R&D Project Selection using the analytic network process", *IEEE Trans. Engin. Manage.*, vol. 49, pp. 59–66, 2002.

[12] M. Pipattanasomporn, "A study of remote area internet access with embedded power generation", Ph.D. thesis. Alexandria, University of Virginia, 2004.

[13] T. N. Andrew and D. Petkov, "The need for a systems thinking approach to the planning of rural telecommunications infrastructure", *Telecommun. Policy*, vol. 27, no. 1–2, pp. 75–93, 2003.

[14] T. Saaty, *Decision Making with Dependence and Feedback, the Analytic Network Process*. Pittsburgh: RWS Publications, 1996.

[15] T. L. Saaty, *Decision Making with Dependence and Feedback, the Analytic Network Process*, Pittsburgh: RWS Publications, 2001.

[16] H. Lee, C. Kim, H. Cho, and Y. Park, "An ANP-based technology network for identification of core technologies: a case of telecommunication technologies", *Expert Syst. Appl.* (in press).

[17] R. W. Saaty, *Decision making in complex environments: a tutorial for the SuperDecisions software*. Creative Decisions Foundation, 2003.

[18] C. Douligeris and J. Pereira, "Telecommunications quality study using the analytic hierarchy process", *IEEE J. Selec. Areas Commun.*, vol. 12, no. 2, pp. 241–250 1994.

[19] Y. Kawasumi, Y. Miyoshi, and K. Koizumi, "How to provide broadband access for rural and remote areas of developing countries", in *Proc. PTC'07: Telecom with Vision*, Honolulu, Hawaii, 2007.

[20] T. L. Saaty, *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks*. New York: Springer, 2006.

[21] T. L. Saaty, *The Analytical Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York: McGraw-Hill, 1980.

[22] T. L. Saaty, *Multicriteria Decision Making: The Analytic Hierarchy Process*. Pittsburgh: University of Pittsburgh Press, 1988.

Yousef Gasiea received the B.Eng. degree in communications engineering from the Higher Institute of Electronics, Ben Walid, Libya, in 1988, and the M.Sc. degree in engineering management from Al-Fateh University, Tripoli, Libya, in 1999. He has worked as an electronic and telecommunications engineer, supervising a field study to measure the signal strength received from various geo-stationary satellites, and installation and maintenance of PABX, TVRO and SMATV systems. Currently he works towards his Ph.D. in engineering management at the University of Manchester, UK. His current research interests include planning of rural telecommunications infrastructure, multicriteria decision methods, decision support systems and analytical decision processes.

yousef.gasiea@postgrad.manchester.ac.uk
School of Mechanical, Aerospace and Civil Engineering
University of Manchester
M60 1QD Manchester, United Kingdom

Margaret Emsley is civil and structural engineering graduate whose early career involved working on construction sites in the UK and a 12 month design office secondment. She was then a Research Assistant at Loughborough University, working on productivity measurement, before commencing her Ph.D. which was concerned with modelling the location of tower cranes. For the past 24 years she has worked at the University of Manchester where she is a Senior Lecturer in the School of Mechanical, Aerospace and Civil Engineering. Her research interests are in the field of construction and project management and include the application of neural networks (where she has carried out funded research into their use in the prediction of construction cost), game theory and simulation modelling.

margaret.emsley@manchester.ac.uk
School of Mechanical, Aerospace and Civil Engineering
University of Manchester
M60 1QD Manchester, United Kingdom

Ludmil Mikhailov received the M.Sc. degree in automatic control from the Technical University in Sofia, Bulgaria, and the Ph.D. degree in technical cybernetics from the Bulgarian Higher Certifying Commission, in 1976 and 1982, respectively. He was an Associated Professor in the Institute of Control and System Research at the Bulgarian Academy of Sciences. Currently he is a Senior lecturer at the Manchester Business School, the University of Manchester. He is the author of about 90 technical papers in peer-reviewed journals and international conferences. His current research interests include multiple criteria decision analysis, fuzzy logic systems, decision-making under uncertainty, and intelligent decision support systems.
e-mail: ludi.mikhailov@manchester.ac.uk
Manchester Business School
University of Manchester
M15 6PB Manchester, United Kingdom

Paper

Stochastic Models in Techno-Economic Analysis of Broadband Access Networks

Paweł Olender

Abstract—Development of networks, specially access networks, is very important and urgent task nowadays. However, it turns out that this segment of telecommunication networks is the most expensive and complicated part of this undertaking. Therefore, the thorough analyses are carried out to determine the best solution under specific circumstances before any decisions are made. This paper presents techno-economic model, which was implemented and used to carry out analyses for one of the biggest city in Poland. To take uncertainty into consideration the stochastic approach was applied providing more robust solution, therefore, improving the safety of investment. Analyses concern FTTH (fibre to the home) technology, type of generic FTTx network architecture. It uses optical fibre in local telecommunication loop, what is becoming more and more popular. Presented results show the usefulness of techno-economic surveys in planning access networks development. The appropriate choice of network parameters, such as the aggregation ratio, is essential and could significantly influence the investment profitability.

Keywords—broadband, decision support, FTTH, network planning, stochastic programming.

1. Introduction

Demand for network bandwidth has been rapidly growing for several years and telecommunication networks become basic information exchange channel for information society and knowledge-based economy. If majority of backbone network has been upgraded already, then access networks are usually based on old infrastructure and are a bottleneck in subscriber connections. Therefore, development of this segment became crucial in improving broadband accessibility. On the other hand, however, it is the most expensive network layer, hence detailed analyses are carried out before any decisions are made. It allows decision-makers to find out possible future results of network development venture. To perform such surveys in reasonable time and comparable way it is necessary to use dedicated tools. Projects were conducted to develop a methodology for such analyses. One of them was Tool for Introduction Scenario and Techno-Economic Evaluation of Access Network (TITAN) executed within the Research in Advance Communications in Europe Programme (RACE). It addressed different aspects of the access networks' development: geometrical models for the distribution, evolution of component price in time or operational cost approach [1]. The methodology was implemented basing on Excel spreadsheet software. Excel was also used in implementation of worldwide interoperability for microwave access (WiMAX) model created in National Institute of Telecommunications (Warsaw). This model consists of 3 parts: demand, technological and economic. It is thoroughly described in [2]. There were also different cost analyses, concerning fiber networks either [3], [4].

Observing situation on the market it seems there is still a need to develop such applications for open access, for example, in case of public tender to verify and compare received proposals. Private companies do not usually make their tools available or it is very expensive. They would rather offer execution of the analysis and provide results. There is a risk of subjectivity. For example, some equipment can be favoured in respect of marketing connections.

Surveys show the nature of network planning process is strongly affected by uncertainty. Presented idea tries taking this aspects into consideration within analysis. Inclusion of random elements in techno-economic model ensures reflecting real circumstances in a better way and leads to more robust solution. Analysis of different scenarios is also helpful in making the decision maker aware of potential threats. Furthermore, mentioned research are mainly based on simulation. Here an optimization approach is partly suggested to obtain the best possible solution.

As network technology fibre to the home (FTTH) was assumed, which meets all demands of the newest services (video on demand, video conference, online games, etc.). Although it seems the most expensive technology it could turn out the most profitable and safest solution in time perspective. It results from laborious and costly installation of passive telecommunication infrastructure and its long depreciation period. For this reason hybrid fiber/copper network could become investment trap because of insufficient bandwidth in few years, what will forced operators to replace copper cables with fibres [5].

The characteristic of network development process and general conception of uncertainty modeling are described in Section 2. Section 3 is devoted to FTTH technology. Structure of the model and its particular modules are presented in Section 4, and Section 5 discusses experiments that were carried out and their results. Summary and potential future research are outlined in Section 6.

2. Dynamics and Uncertainty

The development of access network is usually a large-scale process and lasts several of years. In view of this specificity presented model takes into consideration long-lasting investment character, where the time period is discrete. For that reason some of future data have to be assumed, what of course may bring inaccuracy. It is even more obvious when one considers what kind of data is used in model. For example, demand which determine how many users will be eager to be connected to the network. It is very difficult to precisely forecast such data with regard to the subjective individual potential subscribers' decisions. This uncertainty should be considered to receive more appropriate results.

To address the above-mentioned issues multistage stochastic programming was used. It is a generalized form of more popular two stage approach, where in first stage decision is made before, and in second stage after an outcome of the random process. The second stage decision can be treated as a response to the realization of the stochastic process aimed at minimizing possible losses, caused by first stage decision. Multistage problem is similar, but the outcomes are revealed and decisions are made sequentially (Fig. 1). The general explanation of stochastic programming is in [6]. More detailed and extensive description can be found in [7], [8].

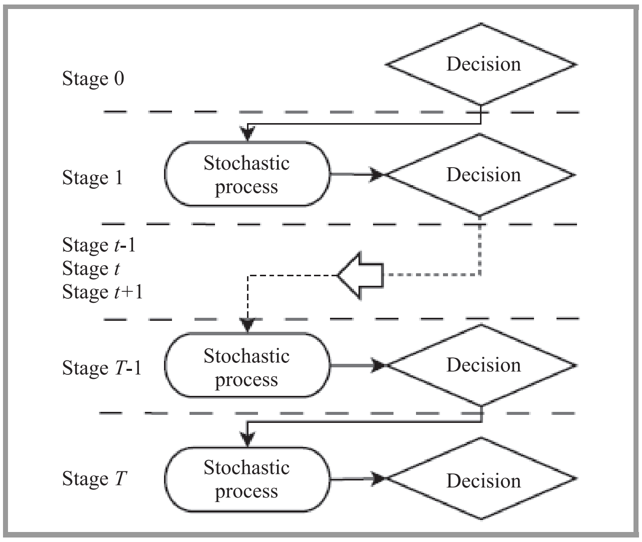


Fig. 1. Decision process in multistage stochastic programming.

As random variable, in the presented model, the take up rate was assumed (which, as mentioned before, seems the most unpredictable). The stochastic process was expressed as a finite set of different scenarios. To formulate the problem as typical mathematical programming a scenario tree was applied. It is a standard tree with a single root, where each node has unique predecessor (except the root) and possibly several successors (leaves have not successors). The nodes at the same level correspond to various scenarios in one period of time. Assuming that the root corresponds to 0 period and venture lasts T periods, there are $T+1$ levels in the scenario tree.

Construction of scenario tree. In the presented model transition-based approach is used to build the scenario tree. This method was described in detail in [9], so here main concept will be explained and a way to use it in assessing the demand.

This approach is based on transitions (branches) between nodes from different levels. A number of branches and their probability are specified to determine the possible realizations of the stochastic process. Starting from the root node this information allows to define the nodes at further levels. To simplify, a symmetrical tree was assumed, what results in the same number and probability distribution of transitions for the nodes at the same level.

To formally define the transitions between t and $t+1$ stages the following functions are defined:

- $f(t)$: the number of transitions from each node at stage t;
- $p(t,i)$: the probability of transition $i \in \{1,\ldots,f(t)\}$ from each node at stage t;
- $d(t,i)$: the change of take up rate corresponding to transition $i \in \{1,\ldots,f(t)\}$ at stage t.

Giving the above values explicit for $0 \leq t < T$ and $i \in \{1,\ldots,f(t)\}$ is sufficient to build the scenario tree.

Each node can be unambiguously determine by a pair (t,n), where t is a number of the stage and n is a number of the node in this particular stage. Moreover, because any node has a unique predecessor it is possible to determine a whole path from the root node to the specific (t,n) node. To navigate through the tree some auxiliary functions have to be defined, like a number of the nodes in the given period, predecessor function, which for (t,n) node gives the number of the predecessor node in $t-1$ period or a number of the last transition from the predecessor to the actual node. To give some idea about the nature of these functions, one of them is explained below. Let us consider the first mentioned function, the number of the nodes in the given period. According to prior assumption that the scenario tree is symmetrical and its consequences, the number of the nodes in period t equals a product of $f(t-1)$ and the number of the nodes in period $t-1$. Bearing in mind that there is single root node the function can be defined as

$$N(t) = \begin{cases} 1 & \text{for} \quad t=0 \\ N(t-1)f(t-1) & \text{for} \quad 0<t\leq T. \end{cases} \tag{1}$$

Finally, there are two the most important values corresponding to each node: take up rate $D(t,n)$ and unconditional probability $P(t,n)$. $D(t,n)$ is a possible value of total take up rate, which can occur with $P(t,n)$ probability in period t. $D(t,n)$ is a simple sum of the changes of take up rate $d(t,i)$, which belong to the path from the root to node (t,n) and initial take up rate in period 0. The probability $P(t,n)$ can be computed by multiplication the probability of predecessor and the probability of transition, which leads from predecessor to current node (t,n). These two parameters determine all possible scenarios in a given time horizon.

Of course in reality, construction of accurate scenario tree, or more general demand model, which takes into consideration all possible scenarios, is hard problem itself. Therefore, author is aware of some simplifications in the above scenario representation. However, presented approach allows to model in a simplified way different demand forecast evolutions, both total in whole time horizon and dynamics in particular stages.

3. The Outline of FTTH Technology

Fibre technology is getting more and more popular in access networks. There are various types of FTTx networks depending on fibre saturation of connection between central office and subscriber [10]. FTTH is characterized by bringing the fibre medium directly to the end-user and for this reason is considered to be the most future-proof access technology and hence also as the target for evolutionary path of access networks. This expectation is supported by unlimited bandwidth of passive infrastructure, because bandwidth of such network is limited only by active equipment like transmitters and receivers [11].

Types of architectures. Two main different FTTH architectures can be distinguished: point-to-point and star. In point-to-point architecture there is a dedicated fibre to each subscriber, which connects him directly to the central office. It is the simplest FTTH network to design and maintain. Star architecture, on the other hand, is characterized by sharing one fibre by many customers through a remote node, which is located between subscriber's household and the central office. This node aggregates/splits network traffic from/to different customers. Star architecture can be active, when the remote node is powered, or passive, in other case. Further, the passive star can be divided into single wavelength system (all subscribers use a common wavelength) and wavelength division multiplexed (WDM) system (each subscriber uses a different wavelength).

Main components of FTTH connection are: optical line termination (OLT), optical network unit (ONU) and obviously optical fibres. OLT is a device with optical transceiver and it is usually located in the central office. In point-to-point architecture one OLT port is dedicated to single subscriber, when in star case it is used by group of subscribers. ONU, called also optical network termination (ONT), is situated on the customer premises. This device converts optical signal from the central office to subscriber into its electronic form and electronic signal into optical in the opposite direction. Optical path is composed of fibres and is usually divided into 3 sections:

– feeder cable: first section of connection from the office center, ends in the remote node in the star network;

– distribution cable: links the remote node with network access point;

– drop cable: direct connection between the subscriber and the access point.

In point-to-point network these are all main components and division of the path is obviously quite artificial, because of dedicated fibre to each customer.

The star architecture apart from above mentioned components needs also a few more devices. Mainly it concerns the remote node equipment and specially switches/splitters. In the active star case in the aggregation node there are switches and temperature stabilization system, therefore, the node have to be powered. In the passive star in the remote node splitters are used to distribute optical signal among subscribers.

Presented model assumes the passive star architecture. It is the most popular FTTH architecture today what may result from no need to power the remote nodes and therefore, economic character of such network. Furthermore, the star network has more complex structure in view of feeder cable sharing and for that reason is more complicated to model in comparison to dedicated architecture.

4. Model Specification

4.1. General Structure

Analysis of research, which were carried out so far, led to some assumptions, which had a significant influence on presented model structure. Apart from dealing with uncertainty there were also two others important premises: optimization character and linearity. The former assures the best solution in specific circumstances, in contrast to simulation, which only gives answers to fixed inputs, causing searching process of satisfactory solution very arduous. What even more important it is not known if calculated solution can be improved. The latter was assumed in order to simplify computation, because stochastic problems are usually quite big, and to use freely available software to solve linear and integer programming problems.

Based on the above assumptions, to reflect characteristic of the passive optical network (PON), a compromise was

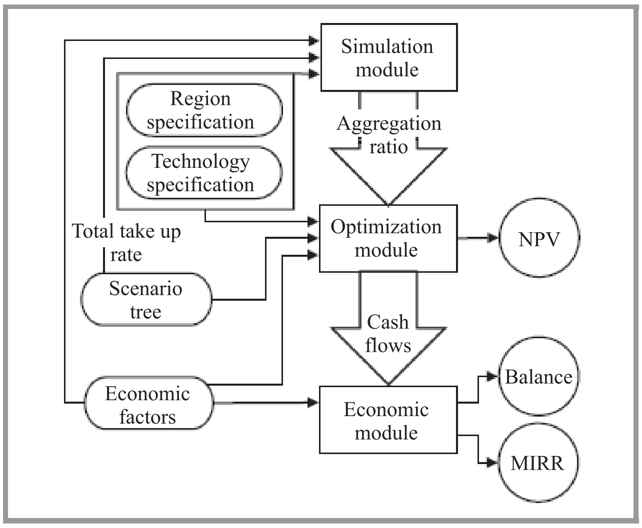


Fig. 2. General structure of model.

made, which resulted in module-based architecture of the model (Fig. 2). There are 3 basic modules:

– nonlinear simulation module: calculates the optimal aggregation rate in the remote node;

– optimization module: computes how many and which customers should be connected to maximize objective function (net present value – NPV, was chosen), using optimal aggregation rate;

– economic module: the simplest part, which calculates additional economic parameters helpful in assessing investments, using cash flows from optimization module.

4.2. Network Dimensioning

Before particular modules are described some information will be given about geometric model used for network dimensioning. One should realize that cost of passive infrastructure, specially for cable networks, constitutes a quite big part of total costs, specially when trenches have to be dug. Thus it is important to determine basic measurements of access network.

Presented model has a hierarchical structure with four levels (to simplify square areas on each level were assumed) and is an extended approach, which has been proposed in [3]:

– access region,

– access area,

– aggregation area,

– connection zone.

The access region is a whole geographic region, where access network is built. Roughly in the middle there is the central office connected to the backbone network. The access region is composed of the access areas, which in reality correspond to districts, estates or small rural regions and are characterized by a given area, household density, distance to the central office and numbers of different kinds of potential subscribers. Uniform distribution of subscribers density within individual kind of the customers in the access areas was assumed.

Each access area is divided into the same size aggregation areas exact to multiple size of this aggregation area. In the middle there is the remote node which aggregates network traffic from subscribers and links them to the central office. Next the aggregation area is divided into connection zones, where single building with potential subscribers was assumed. In the building there might be different number of the subscribers depending on their kind and type of access area (precisely on its population density).

According to above it is possible to determine basic measurements of the network. Trenches schema is shown in Fig. 3(a). When u stands for length of connection zone side, a for length of aggregation area side (in u unit)

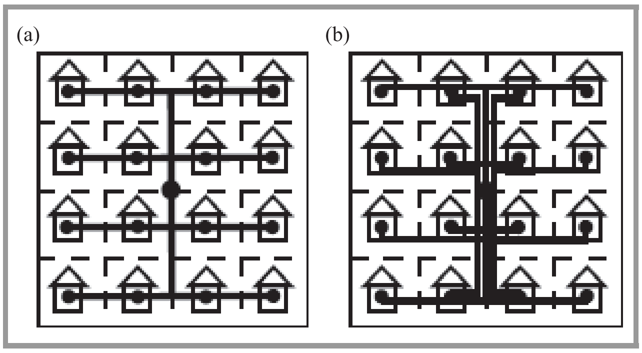


Fig. 3. Schema of aggregation area: (a) trenches; (b) fibres.

and n is a number of the buildings in this aggregation area, length of the trenches L_D can be calculated as

$$L_D = (n-1)\,u, \tag{2}$$

where $n = a^2$. Next, for specific network topology, i.e., star topology (Fig. 3(b)), length of the fibre used to connect customers to the remote node can be calculated. Assuming initially single fibre to each building and even value of a it could be done as

$$L_f = \left(n^{\frac{3}{2}}/2\right)u. \tag{3}$$

Keeping in mind that in one building there can be several customers, above simple dependency should be modified to take into account bringing several fibres to single building. Denoting set of customers types as K, it is possible to express total number of customers in the given access area as $C = \sum_{k\in K} c_k$. Assuming g_k as an average number of k-type customers per building, total number of buildings can be calculated as $B = \sum_{k\in K} b_k$, where b_k represents a number of buildings occupied by k-type customers $b_k = c_k/g_k$. Further, determining c_a as a number of customers in the aggregation area the number of buildings in the aggregation area b_a can be calculated as $b_a = \frac{c_a}{C}B$. These lead to complete formula determining length of the fibre in the aggregation area:

$$L_f = \frac{(b_a)^{\frac{3}{2}}}{2}u\left(\sum_{k\in K}\frac{b_k}{B}g_k\right), \tag{4}$$

what can be simplified to

$$L_f = \frac{(b_a)^{\frac{3}{2}}}{2}u\left(\sum_{k\in K}\frac{c_k}{B}\right), \tag{5}$$

where $u = \sqrt{\frac{A}{B}}$ and A is an area of the access area.

4.3. Simulation Module

The aim of the simulation is to calculate the optimal aggregation rate in the remote node. It is aggregation rate which minimize expected infrastructure cost for the given scenario tree in the access area. In other words it has to be decided how many subscribers should be connected to one aggregation node, therefore, how many splitters should be installed. Because this decision has to be made ultimately at the beginning of investment but it should minimize final network, the total take up rate in different scenarios at the last stage are considered only.

As it was considered in [4], there are two opposite aims which determine the optimal aggregation ratio. It results from the cost of optical fibres, splitters and OLT ports. On the one hand the less splitters in the remote node, the smaller aggregation area, therefore shorter distance from the node to the subscriber and lower cost of fibres. On the other hand, however, the less splitters in the remote node, the greater number of nodes and the higher probability of using greater total number of splitters in the access area. Obviously it is true only for take up rate less than 100%. To better understand the second dependency one can imagine small access area which is divided into two aggregation areas, each with the remote node with single splitter. In such situation each splitter has to be installed even if there are only two subscribers, each in different aggregation area. If these two remote nodes are replaced by one node with two splitters, the second splitter will be used only if there are at least $r+1$ subscribers, where r is the split ratio. The number of splitters determines also the number of OLT port in the central office, thus higher aggregation ratio decreases cost of splitters and OLT ports.

Assuming m as the aggregation ratio (it means that there would be installed m splitters in the remote node if take up rate was 100%) and r as the split ratio, there will be mr customers in the aggregation area. Assuming further binomial distribution of customers take up rate, what means each customer takes a decision independently, keeping in mind uniform distribution of customers density, and using dependency (Eq. 5), length of the fibre can be calculated as

$$L_{f_{m,n}} = \frac{1}{2}\left(\frac{mr}{C}\right)^{\frac{3}{2}} \sqrt{A}\left(\sum_{k\in K} c_k h_{k,n}\right), \tag{6}$$

where $h_{k,n}$ is a fraction of k-type customers, which want to be connected to the network, when n scenario occurs.

It is left to calculate the real number of splitters, which are used in the remote node, at $h_{k,n}$ take up rate for $k \in K$ and n scenario. Because there are mr customers in the aggregation area and fraction of k-type customers equals c_k/C, the number of subscribers is

$$N_{ConSub_{m,n}} = mr\left(\sum_{k\in K} \frac{c_k}{C} h_{k,n}\right). \tag{7}$$

Hence keeping in mind r as the split ratio, the number of splitters in the remote node can be expressed as

$$N_{spl_{m,n}} = \left\lceil m\left(\sum_{k\in K} \frac{c_k}{C} h_{k,n}\right)\right\rceil, \tag{8}$$

where $\lceil x \rceil$ is the smallest integer not less than x. To be precisely it should be add there will be $\lceil \frac{C}{mr} \rceil$ aggregation areas, when the aggregation ratio is m.

To find the optimal aggregation ratio the average infrastructure cost per connected subscriber is calculated. Assuming that C_{OLT}, C_{spl}, C_f are, respectively, costs of OLT port, splitter and fiber (length unit, e.g., 1 km) this can be done as

$$C_{m,n} = \frac{\left(C_{OLT} + C_{spl}\right) N_{spl_{m,n}} + C_f L_{f_{m,n}}}{N_{ConSub_{m,n}}}. \tag{9}$$

Calculating this cost for each scenario n in the last period, it is possible to determine its expected value for the given scenario tree and aggregation ratio m:

$$E\left(C_m\right) = \sum_n C_{m,n} P_n, \tag{10}$$

where P_n is probability of scenario n.

Assuming $M_{\max}$ as maximal number of splitters that can be installed in the remote node, module computes expected cost $E\left(C_m\right)$ for aggregation ratio $1 \le m \le M_{\max}$ and returns one which minimizes it.

4.4. Optimization Module

The most important and complicated part of the presented model is optimization module, which merges technological and economic aspects. All details of this module was described thoroughly in [12], thus here some basic assumptions and general structure will be only explained.

Module on the basis of given demand calculates the necessary amount of equipment, its cost and possible revenue. Users demand is estimated using the scenario tree, which was described in Section 2. Further, the geometric model and the optimal aggregation rate value from simulation module are used. As decision variables the number of connected and non-connected customers of each type were assumed. The objective function is net present value at the end of assumed time horizon

$$NPV = \sum_t \frac{CF\left(t\right)}{\left(1+r\right)^t}, \tag{11}$$

where r is the depreciation value and $CF(t)$ is the cash flow in period t (all costs and revenues between t and $t+1$).

For the sake of linear assumption, what was mentioned earlier in the model structure description, amount of all network components, which have nonlinear dependency on the number of subscribers, are calculated to satisfy total demand (to connect all eager customers). Model takes into account both, capital (CAPEX) and operational (OPEX) expenditure. CAPEX includes costs of:

- trenches and ducts,
- fibres,
- central office,
- remote nodes,
- customer premises equipment.

OPEX is calculated proportional to CAPEX, according to classes suggested in TITAN methodology [1]. Revenue comes from customers connection and monthly fees.

4.5. Economic Module

In third module selected economic factors are computed on the basis of cash flows from optimization module. Presented model takes into consideration two additional values:

- balance,
- modified internal rate of return (MIRR).

Balance is a simple sum of revenues and costs. Due to this simplicity it is a very common economic factor to assess various ventures. It could be treated as down limitation of investment profitability – if balance is less than zero, all values, which take into consideration discount aspect (i.e., NPV), will also show unprofitability of such undertaking (typical investment is the most costly at the beginning). Moreover, the balance is objective in contrast to discounted values, which depend on depreciation rate. For these reasons it is probably the most universal measure to compare different investments, specially those with similar time horizon.
The second measure is modified internal rate of return, which as IRR is relative value of venture efficiency, unlike absolute values as balance or NPV, which reflect size differences between various investments. MIRR in contrast to IRR assumes the positive cash flows from a project are reinvested at the rate equates usually to cost of capital (depreciation rate used in NPV calculation). In case of IRR the positive cash flows are reinvested at the internal return rate. For that reason MIRR more accurately reflects the nature of access network development project. As it was described before development of network is a long-lasting undertaking, specific rather repeatable, what explains use of different reinvested rate instead of common internal return rate.
In other words MIRR is the deprecation rate for which discounted total future value of all positive cash flows equals total present value of all negative cash flows reached from investment. Formally this can be expressed as

$$\sum_{t=0}^{T} \frac{CFI_t}{(1+r)^t} = \frac{\sum_{t=0}^{T} CFO_t\,(1+r)^{T-t}}{(1+MIRR)^T} \qquad (12)$$

and therefore

$$MIRR = \sqrt[T]{\frac{\sum_{t=0}^{T} CFO_t\,(1+r)^{T-t}}{\sum_{t=0}^{T} \frac{CFI_t}{(1+r)^t}}} - 1, \qquad (13)$$

where CFI_t is the negative cash flow in period t, CFO_t is the positive cash flow in period t, r is the depreciation rate (cost of capital) and T is the time horizon of investment.
As it can be seen MIRR is not only more appropriate to network development investment in due to its specificity but is also simpler to compute in comparison to IRR, for which it is necessary to use reverse-simulation method.
Decision is made by analogy with IRR, therefore the investment is profitable when MIRR is greater than cost of capital and obviously the greater it is, the better.

5. Techno-Economic Analysis

Described model was applied to perform techno-economic analysis for the city Łódź (Poland) and surroundings. The region was chosen with regard to available data, which were obtained thanks to [2] authors' kindness. That study also concerns development of the access network but for WiMAX technology, thus it was possible to compare some of the results in order to general verification. Moreover, the region of Łódź varies widely in population density, therefore it is good for analysis. Three types of the access areas were distinguished: urban, suburban and rural.
Time horizon of the investment was assumed to be 6 years. It was subjective decision on the basis of telecommunications market profile. On the one hand it should be horizon long enough to develop the access network but on the other hand short enough to make economic and demand forecasts reliable as much as possible and to convince investors to wait for profits. All presented below experiments were carried out using the hypothetical scenario tree with 24 scenarios in the last period, where take up rate varies from 10% to 22% for individual and business customers (public customers like schools and municipality offices were connected in 100%). This simple estimation was a result of study various demand forecasts and experience from FTTH project in other countries, which were obtained mainly from the Internet.

5.1. Aggregation Ratio

Results from the simulation module will be discussed at first. Optimal aggregation ratio for 3 selected access areas with different population density were calculated. For each access area expected cost was computed for 3 different demand cases:

- set of possible scenarios with their probability in the last period;
- the lowest take up rate (pessimistic scenario);
- the highest take up rate (optimistic scenario).

In the second and third cases there was assumed only one possible scenario with 100% probability. In Fig. 4 results for low population density area are illustrated.

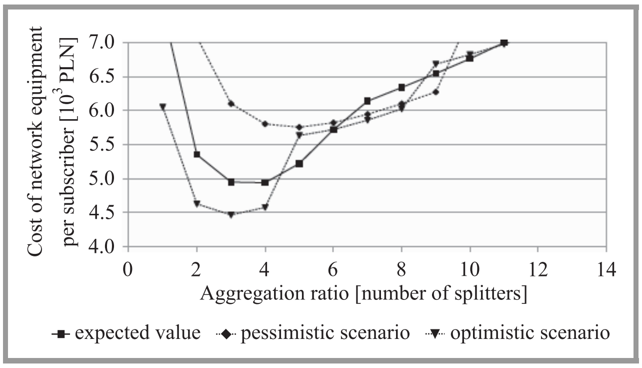


Fig. 4. Simulation module results for low population density area.

First, one can see that single splitter in the remote node is the worst solution, giving significant higher costs than higher aggregation ratio in all aggregation areas, although it could seem intuitively that installing splitters the closest as possible to the subscribers to reduce the use of fibres would be the best approach.
Further, it turns out that generally the aggregation ratio for pessimistic scenario is greater than its equivalent for optimistic one. Therefore, it seems the lower the take up rate, the higher aggregation ratio is more profitable. Because the aggregation ratio is increasing with the reduction of the subscribers number, this means the costs of fibres have less impact than the costs in the remote nodes and central office. It is different when changes of the population density are considered. In rural areas the optimal aggregation ratio is lower than in areas with greater population density. Therefore, one can come to a conclusion the main factor, which affects the costs this time, is the distance between the subscribers and the remote node.
Moreover, growth of population density, as well as take up rate, decreases the costs per subscriber, thus the higher payback from investment can be expected. To sum up it seems FTTH access network investment is positive correlated with the number of subscribers.

5.2. Economic Analysis

On the basis of optimal aggregation ratio from simulation module further experiments were carried out. Some economic values were calculated to assess investment profitability for various cases.
At the beginning different population density access areas were compared (Fig. 5). As it was mentioned before 3 types of areas were considered: low population density rural area, average population density suburban area and high population density urban area. As it can be seen in Table 1 only urban area is economically profitable, what is confirmed by both NPV greater than 0 and MIRR greater than assumed 15% discount rate (cost of capital). However, balance is positive for all areas what can mean that in longer time horizon each of these investments can be profitable, especially as maintain costs of fibre network are lower in comparison to copper one.
It is also interesting to compare suburban and rural areas. Although NPV is about 2.5 times lower for suburban area, MIRR are very similar and balance is much greater than for rural area. It results mainly from different scale of investment, because in suburban area initial expenditures are much higher. For this reason it seems that for longer time horizon the suburban area will be more profitable.

All these observations lead to the conclusion that investment in FTTH access network is positive correlated with the number and density of subscribers.
Similar analysis was carried out for the case when all interested customers were connected to the access network (additional constraint on the unsatisfied demand). Such requirement seems rational for example when a public organization subsidizes access network investment in non-

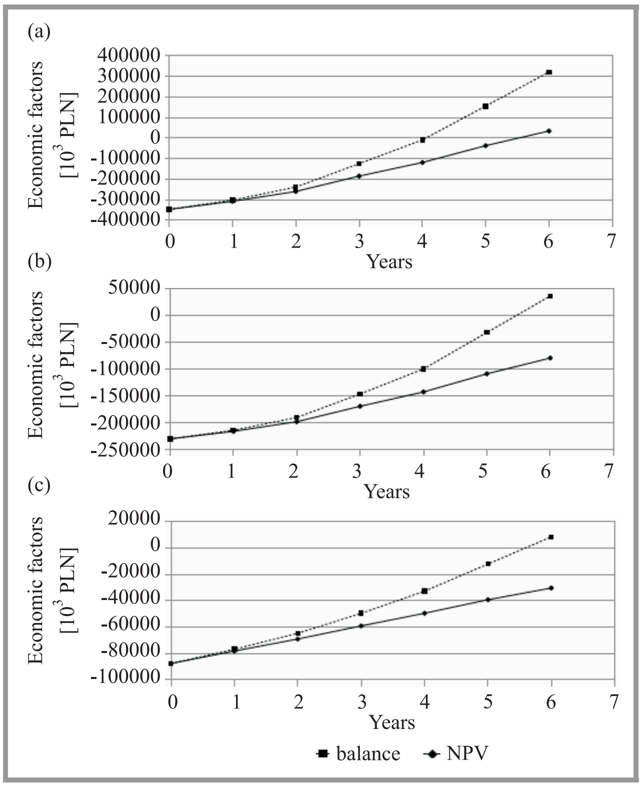


Fig. 5. Economic factors for different population density areas: (a) high density; (b) medium density; (c) low density.

profitable areas. In this particular situation such assumption aggravated economic results only slightly. It also allowed to estimate the value of subsidy required to make investment profitable in suburban and rural areas. It equals about 80 and 30 millions PLN, respectively.

Table 1
Economic factors for different population density areas

Area	Balance [10^3 PLN]	NPV [10^3 PLN]	MIRR [%]
Rural	8481.9	–30347.8	7.14
Suburban	35205.7	–80540.6	7.08
Urban	319654.7	33828.8	16.80

Another experiment concerned duct system. Because trenching primary duct system is very expensive and the city Łódź keeps control of the sewage system, the possibility to use it was considered to reduce capital expenditure. It was assumed the sewage system exists in suburban and urban areas. As it is shown in Fig. 6 this approach improves significantly efficiency of investments and makes the access network development profitable also in medium population density area. Payback period (time for which NPV = 0) is about 4 years long in suburban area and shortens to about 3.5 years in urban area.
The City Hall of Łódź had a clear vision of network ownership and management, where the city was an owner of the network infrastructure but direct customers services

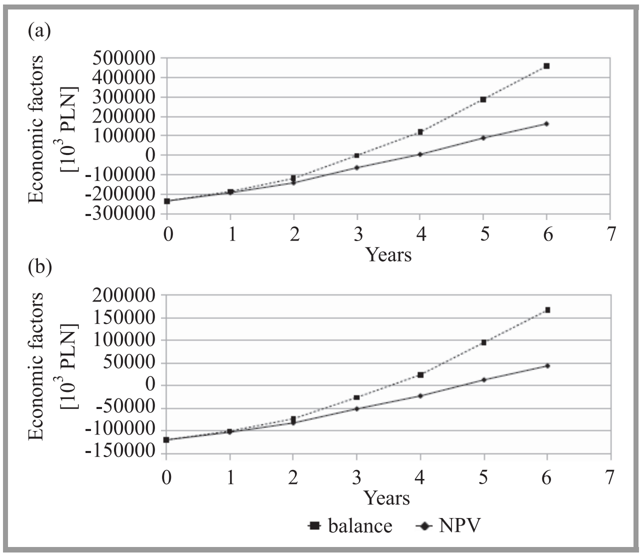


Fig. 6. Economic factors using the sewage system: (a) high density; (b) medium density.

were offered by the service operators [2]. Due to this business model of investment, there was a need to decide how to share income between the network owner and the service providers. In previous experiments it was assumed 20% of income goes to the service operators. Here (Table 2) results for two other values, 35% and 50%, are presented. In case of 35% value investment is still relatively profitable, but when profit margin reaches 50% undertaken will cross the profitability limit. Presented results concern high population density area with the primary duct system.

Table 2
Economic factors for high population density area and different service operators profit margin

Profit margin [%]	Balance [10^3 PLN]	NPV [10^3 PLN]	MIRR [%]
35	269781.1	62105.1	19.64
50	101031.6	–21429.0	12.64

Income division ratio plays an important role in business model, thus in reality it should be subject of thorough negotiations, taking into account various aspects, for example, competition.

6. Conclusions and Future Work

Presented results have shown the usefulness of techno-economic analyses in planning access networks development. The appropriate choice of network parameters, such as the aggregation ratio, is essential and could significantly influence the investment profitability. Additionally stochastic approach, which takes uncertainty into consideration, provides for a more robust solution due to changes in the volume of customer demand. Suggested decisions are optimal in sense of expected value of NPV. Such method improves the safety of investment as opposed to the formulation in which random variables are replaced by expected values what may lead to infeasible solution. Analysis of different scenarios can also make the decision maker aware of potential threats.

One of the biggest problem in creating the model was limited availability of specific data. It caused some simplifications and common sense assumptions. As it turned out, detailed demographic data were not freely available and as it was mentioned before, they were provided by [2] authors. Information about the cost of components was hard to obtain as well, what could be a consequence of quite new FTTH technology.

The most important findings from experiments are not absolute exact value of individual factors, which in view of assumptions and difficulties could be questioned, but the general dependencies and characteristics follow from received results. From economic point of view, it seems the fibre access networks are more appropriate, when there is a significant number of subscribers. It results mainly from high initial investments, which are necessary to provide required infrastructure to cover the given access area, and this is related to costly civil work and trenching. For this reason, the FTTH network is unprofitable in low population density area and needs subsidy under such circumstances. On the other hand, there is an interesting opportunity for the areas where there is some kind of ducts system which can be easily adopted to distribute fibre cables. In this case, the development of the FTTH network is much less expensive, shortening significantly the payback period. From decision maker point of view, such property is also a guidance to develop the passive infrastructure within other investments like road construction, although one has to realize it requires a long-term planning period.

Apart from data availability difficulties it is necessary to mention other shortcomings and possible improvements to deal with. First is the computational complexity. One should realize the number of scenarios grows exponentially with the number of periods. Presented experiments were carried out using small scenario tree, what in view of small range of take up rate parameter was not the problem, but in general it poses the computational limits when the parameters vary greatly. To overcome this limitation the decomposition method should be considered. Secondly it should be also remembered about some simplifications in the scenario tree building method. It is the important issue which requires further detailed research. Some algorithms to build the scenario tree should be work out to ensure the most appropriate representation of the possible trends of events.

The issue of robustness in such applications may be also an interesting subject of further studies. It concerns, e.g., the choice of objectives. Some measure of risk could be applied as decision maker might be interested in the variability of returns associated with a plan. It may lead to multicriteria optimization and more complex utility function.

References

[1] B. T. Olsen et al., "Techno-economic evaluation of narrowband and broadband access network alternatives and evolution scenario assessment", *IEEE J. Selec. Areas in Commun.*, vol. 14, no. 6, pp. 1184–1203, 1996.

[2] J. Granat, P. Białoń, A. Gosk, and M. Salwa, "Zarządzanie usługami informatycznymi oraz infrastrukturą informatyczną". Część B: "Modele wspierające rozwój regionalnych sieci dostępowych", Rep. 06.30.002.6. Warsaw: National Institute of Telecommunications, 2006 (in Polish).

[3] H. Sigurdsson, M. Falch, J.-C. Point, and P. Million, "BREAD Deliverable 2.4–3.4: Techno-economic study", 2006-05-10 [Online]. Available: http://www.viskan.net/publications/BREAD-DeliverableD24-D34-draft.pdf

[4] I. Chlamtac, A. Gumaste, and C. A. Szabo, *Broadband services: Business models and technologies for community networks*. Chichester: Wiley, 2005.

[5] K. Borzycki, "Światłowodowe sieci dostępowe", *Telekomunikacja i Techniki Informacyjne*, no. 1–2, pp. 71–99, 2008 (in Polish).

[6] S. Sen and J. L. Higle, "An introductory tutorial on stochastic linear programming models", *Interfaces*, vol. 29, no. 2, pp. 33–61, 1999.

[7] Y. Ermoliev, R. J.-B. Wets, *Numerical Techniques for Stochastic Optimization*. Berlin: Springer-Verlag, 1988.

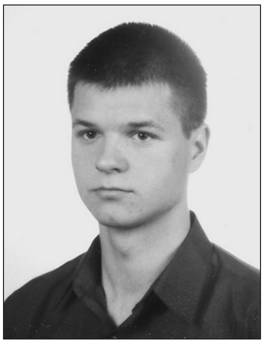
[8] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, New York: Springer, 1999.

[9] J. Thenie, C. van Delft. and J.-P. Vial, "Automatic formulation of stochastic programs via an algebraic modeling language", *Comput. Manage. Sci.*, vol. 4, no. 1, pp. 17–40, 2007.

[10] "Digital sections and digital line system – Optical line systems for local and access networks: Broadband optical access systems based on Passive Optical Networks (PON)", ITU-T Rec. G.983.1 (01/2005).

[11] M. Kunigonis, "FTTH explained: delivering efficient customer bandwidth and enhanced services" [Online]. Available: http://www.iec.org/online/tutorials/

[12] M. Pajer, P. Celej, A. Gosk, J. Granat, and P. Olender, "Zorientowane usługowo rozproszone systemy dostarczania informacji decyzyjnych – rozwój wybranych serwisów". Część B: "Projekt budowy systemu wspomagania decyzji w rozwoju dostępowych sieci szerokopasmowych w technologii FTTH według paradygmatu SOA", Rep. 06.30.003.8. Warsaw: National Institute of Telecommunications, 2008 (in Polish).



Paweł Olender received the M.Sc. in computer science from the Warsaw University of Technology, Poland, in 2008. Currently he is a Ph.D. student in computer science at the Institute of Control and Computation Engineering at the Warsaw University of Technology. He is employed by the National Institute of Telecommunications in Warsaw. Has participated in projects related to data warehousing and analysis for a telecommunication operator. His research interests are focused on modeling, decision support, optimization, data mining and multi-agent systems.

e-mail: P.Olender@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

e-mail: P.Olender@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Paper

A Framework for Evaluation of Communication Bandwidth Market Models

Wojciech Stańczuk, Piotr Pałka, Józef Lubacz, and Eugeniusz Toczyłowski

Abstract—The article presents a method of analysis of market-based models for resource allocation in communication networks. It consists of several stages: classification of a market model, generation of input data, data adaptation to a tested model, test calculations and, finally, presentation and interpretation of results. A set of general criteria to assess various models has been proposed. Tests are run using dedicated computer applications, data is stored in open XML-based format originated in the multicommodity market model. Network topologies are derived from the SNDlib library.

Keywords*—*auctions, bandwidth trading, resource allocation.

1. Introduction

A dominant form of trading on the market of resources in telecommunications networks is bilateral contracting. Time of negotiations is undesirably long in relation to high dynamics of business processes in the telecommunications market. In addition, the bilateral nature of negotiations reduces the transparency of trade rules. This often enables network operators or service providers with significant market power to obtain better trading conditions than it is justified. Therefore, the research is conducted on innovative mechanisms for trading of transport resources in networks to enhance the efficiency of their usage and the quality of conditions of competing for them, particularly in the form of auctions and exchanges. Potential benefits of the introduction of such multilateral trading patterns on the market for telecommunications network capacity are discussed in [1], problems of the organization of network bandwidth exchanges are presented in [2].

The variety of possible models and their variants creates a problem of objective evaluation and the feasibility of mutual comparison. The models may implement many different resource allocation algorithms and apply multiple optimization criteria taking into account economic and technical constraints. The chances of a simple quantitative assessment of one model in comparison to others and indicating its advantages and disadvantages are hindered. This paper is an attempt to develop a methodological approach to testing and comparing models of market-based allocation of capacity in communication networks. The method can be helpful in choosing trade models adequate to specific markets segments in the telecommunications sector.

The structure of the article is as follows: Section 2 briefly describes the survivable network design library (SNDlib) and the multicommodity market model (M^3). Section 3 shows the successive stages of testing. Section 4 presents a set of comparative criteria. Section 5 shows an example of the application of the proposed framework for a selected bandwidth trading model. Section 6 summarizes the results of the research.

2. A Method of Analysis of Market-Based Models

The proposed approach to evaluating market models results in a multi-stage framework. The stages, shown in Fig. 1, are as follows:

- model classification,
- test data generation,
- data adaptation for a model,
- running tests,
- output data analysis.

Each step can be performed independently, using separate tools. Data passed between the successive stages is stored in text files saved in a extensible markup language (XML)-based format.

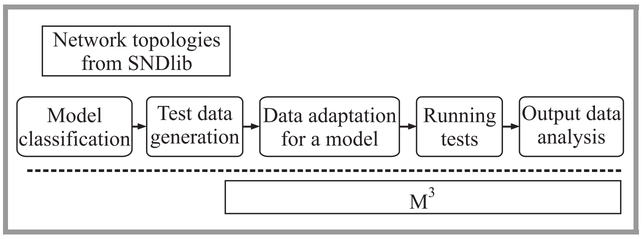


Fig. 1. Schema of the bandwidth model testing framework.

The proposed methodology integrates the results of other research: the data of the SNDlib and the model of a multicommodity market process (M^3).

2.1. The SNDlib Library

The survivable fixed telecommunication network design library [3] is a scientific library sharing exemplary data for problems of the design/dimensioning of communication transport networks.

Examples stored in the library reflect the topology of real networks. All of them are saved in a standardized XML dialect. The main purpose of the library is to collect actual data on research problems and create a platform for exchange of information between scientists and engineers involved in network design. The library comprises network topologies with the structure of network links and bandwidth demands, best solutions and their dual bounds, an up-to-date bibliography and a list of conferences on the subject. The scope of solving methods is broad and includes models of linear/integer programming, branch-and-bound algorithms, column generation, a Lagrangian relaxation and meta-heuristics such as evolutionary algorithms, simulated annealing or taboo search.

2.2. The M^3 Model

Consistent description of a broad scope of potential trade processes on the bandwidth market needs an application of an adequately flexible information model. Such flexibility and a high degree of openness can be obtained by using an information model based on the M^3 [4]. The M^3 is a set of formal models that describe data and communication messages on multi-commodity infrastructure markets. The M^3 model has been adopted in our research for a market of transport resources in telecommunication networks.

The M^3 model enables a generic description of the trading information exchanged between market participants. The data is stored in a special M3-XML dialect allowing the expression of: the existing network infrastructure, the time scale in which the trade is accomplished, the entity structure (sellers, buyers, brokers, leaseholders, etc.), the trade object structure (description of traded goods) and the offers submitted by individual market participants. It contains, in particular, descriptions of market offers: elementary (singlecommodity), integrated (multicommodity), and also grouping: describing more complex relations between elementary commodities or integrated offers with common conditions or resource constraints. The introduction of the mechanism of grouping of the offers facilitates the formulation of non-trivial constraints for each individual market participant, and proper valuation/quoting of the offers.

In practical applications one may use all or just selected elements of the M^3 data model. It is worth noting that the M^3 model and its M3-XML data format do not specify how the trade itself and the allocation of resources are realized.

3. Stages of Preparing and Running Tests

This section presents the sequential stages of preparation and runnig tests of bandwidth market models.

3.1. Classification of the Model

The properties of models originate in economics, game theory principles of mechanism design and in technical features of traded bandwidth. Selecting distinctive features of the models creates a space for their classification and grouping. The set of the classifying features have been developed upon analysis of the network resource structure and the organization of trade processes. The categories of the network design problems from [3] have also been taken into account.

Type of commodities. The number of types of resources in telecommunication networks is significant: there are many technologies used in transmission systems (switching and multipexing techniques). The resources can have a physical or a virtual nature and their detailed specification depends on a layered architecture of modern networks.

Our research is focused on transport communication networks and in this context the general elementary commodity is a bandwidth of a point-to-point connection between a pair of nodes in a specified network layer. It is defined as a network capacity enabling transmission of specified data amount from a source node to a destination node during specified quantum of time.

The bandwidth of a network link can be offered for sell. The bandwidth needed to serve a traffic demand can be purchased – the existence of a direct connection between the pair of nodes of the demand is not necessary, the demand can be realized with one or more paths (consisting of sequences of links).

The trade models can take into account many of characteristic features of bandwidth commodities. The most significant ones are following:

- direction of bandwidth: bandwidth commodities can be directed, undirected (data flow is possible in both directions) or asymmetric (capacity depends on the direction of data flow); bidirectional bandwidth commodities can be modeled with two oppositely directed commodities;
- divisibility of bandwidth: bandwidth can be fully divisible, modular (divisible with the accuracy of units), unit (a particular case of modular bandwidth modeled in combinatorial auctions), predefined (divisible only within specified volumes);
- commodity structure: some trade models omit structural relations between commodities and concern trading only separate elementary commodities; other models enable trading complex structures of commodities, particularly on the demand side: the structures can refer to a set of predefined network paths, a specified set of network links or whole subnetworks (e.g., for purposes of building virtual private networks).

Relations between offers and commodities. The models can take into account different kinds of assignment between traded commodities and related trade offers. One can distinguish:

– elementary offers: a single offer concerns a single elementary commodity;

– bundled offers: a single offer concerns a bundle of elementary commodities, the commodities can be sold/bought in equal or different amounts.

Market participants. The models can describe many organizational forms of trade. The basic division takes into consideration the number and the roles of market participants. One can distinguish:

– single-side models: a market operator sell bandwidth capacity to its clients, or a single client buys services from many competing network operators;

– double-sided models: there are many buyers and sellers.

The important role in trade models is a market operator – the entity that balances the market: allocates resources and sets prices. In centralized models there is one market operator. In distributed models there can be many of them.

Quality of service constraints. Market contracts concerning telecommunication resources usually define a set of parameters describing the quality of service (QoS). The trade models can take those constraints into account, e.g., specifying the maximal length of a path or the maximal delay of a packet.

Resource allocation rule. The network resources are allocated by means of a defined algorithm operating on the available market offers. In the context of transport network resources one can distinguish the basic allocation rules:

– path setting: the required point-to-point bandwidth connections are served with paths consisting of a sequence of communication links;

– allocation of single resources: buy offers concern separate resources that is not bundled without any explicitly expressed relation.

Pricing rule. Market models set contract prices according to a specified rule. Some popular examples of such rules are: English auction, second-price auction, Vickrey-Clarke-Groves auction, dual pricing, etc.

Exchanged messages. The essential feature of a trade model is the type of signals exchanged between market participants and a market operator. The signals can have various forms:

– point characteristics: a traditional market offer indicating commodities, their amounts and offer prices;

– partial characteristics of preferences, e.g., a set of points from utility function or a stepwise offer;

– full characteristics of preferences, e.g., in the form of an utility function.

Market balancing dynamics. The models can be divided according to the time schedule of the market balancing process. One can distinguish two basic classes:

– one-time auctions: market participants submit their offers and then a market operator allocates resources and sets prices taking into account all submitted offers;

– iterated auctions: the final market balance is achieved in a sequence of steps, in which market participants can modify their signals submitted to a market.

Implementation. The models can be implemented with many different techniques and tools originated in operations research, computer programming and mathematics. Examples of the implementation types are following: linear programming, mathematical programming, dynamic programming, parametric equations, heuristics, etc.

3.2. Input Data Generation

One of the purposes of the research presented in this paper is to create a library of test examples for the scientific community involved in the design of market algorithms and models for communication transport networks. Input data for test cases should reflect the size of demand and supply observable in the real network bandwidth market. The test data used for the research may come from the following sources:

– examples of network design problems;

– economic models of supply and demand;

– real data from the telecommunications market.

Trade patterns for transport resources of networks and balancing market offers are conceptually similar to the problems of network design – the relationship between trade mechanisms and designing the network have been discussed in [5]: the demand for network bandwidth and communication links between nodes can be interpreted as a market offer to buy/sell network resources. In order to generate a test network topology and market offers one can therefore use the data from the SNDlib library. The examples from the database cannot be used directly due to their important constraints. Firstly, there is only one link between pairs of nodes in a network topology, while the market models assume the existence of multiple offers on the bandwidth segment between a pair of nodes (a similar requirement applies also to buy offers). Secondly, any pricing information in the examples of the SNDlib is only expressed by determining the cost of installation and expansion of links, while in the trade models one requires the price of bids submitted by buyers and sellers. All the missing elements can be added to the original network design examples, e.g., by means of the pseudo-random generation adjusted to the specific test case.

The second source of the data on the potential offers on the bandwidth market is the use of economic models of supply and demand. Such models in the context of resources in telecommunication networks are under research, especially in the field of modeling network traffic and demand for services. Some analysis of supply and demand can be transferred to the ground of telecommunications from other infrastructure markets such as energy or transportation.

The test data can also come from the analysis of the actual data on transactions accomplished in the real telecommunications market. Acquiring such data is difficult in practice because there are not any network bandwidth exchanges operating on a larger scale, and the information about bilateral contracts between telecommunications companies is generally private and not publicly available. However, there are some internet sites showing examples of bandwidth prices in certain local markets, e.g., the U.S., and reports on global trends in the development of the telecommunications market, such as the work [6].

3.3. Data Adaptation – Conversion to M^3 Format

The adopted information model for the framework is the M^3 – any data of test cases should be saved in the M3-XML format. This can be achieved in two ways.
The first one is the direct generation of input data in the desired format. This requires dedicated computer tools allowing editing or automatic generation of the test data.
The second approach assumes the use of examples contained in the SNDlib library enriched with the entity structure, the object structure and trade offers. Choosing an XML dialect as a data format results in the opportunity to use widely available read/write software libraries for many programming languages, which facilitate the development of new tools for test data generation. The natural way to transform XML data into another format is the use of a extensible stylesheed language transformation (XSLT) [7]. One can also use the query language XQuery [8].
For the purpose of this research a mixed approached has been adopted. A dedicated computer program with a graphical user interface has been designed and implemented. It has been used to enrich the SNDlib examples with lacking elements and save them in the simple dedicated bandwidth market XML (BM-XML) format describing the network nodes and the offers for sale and purchase of network capacity. A series of XSLT transformations processing the BM-XML format into M3-XML format have been used. In this way we have obtained a convenient set of computer tools used for importing and converting examples of network topology to the new library of complete test cases for bandwidth market models.

3.4. Running Tests

The next step is to run required calculations of network capacity allocation and pricing for a tested model with the prepared data. It is assumed that all the input data is stored in the M3-XML format. Therefore any market mechanism, which is a subject to the tests, should be able to read input data and generate output data compatible with the M^3 information model. The flow of data during the execution of the tests is illustrated in Fig. 2.

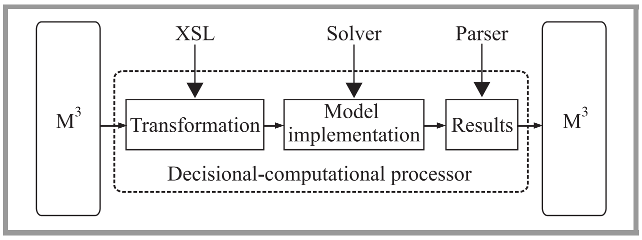


Fig. 2. M3-XML format data flow in running tests.

The concept of a universal decisional-computational processor has been developed [9] and the initial implementation has been made: it solves the tasks described by the data stored in M3-XML format. The corresponding XSLT transformation converts the data to the internal representation of the appropriate model. The processor returns the data in the M3-XML format. Such modular architecture allows independent implementation of the trade models from the evaluation framework: compatibility with the M^3 information model can be achieved by the input and output interfaces. The XSLT transformations can be applied to convert the input data into the format required by the particular computing processor, e.g., a standard mathematical programming solver or a dedicated implementation of the algorithm of a trade model. Examples of such transformations have already been created [4] and they are further extended in our research on new trading models. They convert the data in M3-XML format into GMPL format (gnumath programming language) [10], which can be used by optimization solvers, e.g., GlpSol and AMPL (a mathematical programming language). Linear models can also be easily converted to other LP (linear programming) formats, e.g., the one used by the CPLEX solver.

3.5. Analysis of Output Data

The next step to take after having completed series of computational experiments is to analyze the obtained results. The application of the M^3 information model to describe the output of the tested models facilitates the analysis. If one had different data formats for various trade models the comparison of the results would be much more difficult and time-consuming. The adopted XML-based format allows for simple data conversion to other formats for presentation purposes, such as exporting to spreadsheets or graphical visualization applications.

4. Comparative Criteria

A fundamental part of the evaluation framework is the proper selection of indicators. Trade models are often designed for a specific market context and defined aspects

of resource allocation, so the comparative criteria should be general enough to be applicable to most of them. We propose a set of such indicators, they represent measures originating in economics, game theory and technical efficiency.
A properly designed trade model should strive to meet several desirable properties that make it attractive for a wide range of market participants (traders, decision makers, regulators). These properties can be used as evaluation criteria providing information about the ”quality” of a given model. The set of criteria is divided into three categories: global, individual and technical. The division into the individual and global ones expresses the natural market game between individual participants' interests and global interests. The third category aims in evaluation of the models in terms of their technical efficiency.

4.1. Global Criteria

Economic efficiency. The main measure of economic efficiency is the market surplus. It is defined as the aggregated economic benefits derived from the market exchange of goods. If the market mechanism encourages participants to submit truthful bids, then economic prosperity could be determined accurately using actual bids. If the market offer does not comply with the participant's preference profile, it can be used only as an imprecise measure known as economic benefits. It is the difference between total value of goods purchased and the total value of goods sold on the market, as follows:

$$Q = \sum_{m \in B} d_m e_m - \sum_{l \in S} p_l s_l \,, \tag{1}$$

where e_m is the unit offer price of buy offer m, d_m is realized bandwidth of buy offer m, s_l is the unit offer price of sell offer l, and p_l is realized bandwidth of sell offer l.

Incentive compatibility. The incentive compatibility property holds if no market participant has incentives to report signal different from its type/preferences: no agent has incentives to report an untruthful offer.
Incentive compatible mechanism prevents strategic actions of the participants. The measure of the effectiveness of the mechanism against strategic players is the allocative inefficiency (AI) defined as follows:

$$AI = \frac{Q^0 - Q}{Q^0} 100\% \,, \tag{2}$$

where Q^0 is the economic surplus if every market participant submits a truthful bid, Q is the actual (achieved) value of the economic welfare, where participants can submit bids incompatible with their preferences. It should be noted that the participants may submit bids incompatible with their profile of preferences in order to achieve higher profits.

Budget balance. A trade model has balanced budget if sum of sellers' expenses is equal to sum of buyers' incomes: there is no need to surcharge the trading mechanism, and the mechanism does not earn additional profits.
The first quantitative measure describing this criterion is the value of the difference between total sellers' income SI and total buyers' expense BE related to the total market value. This measure has got the term budget imbalance in relation to total turnover ($RBIT$):

$$RBIT = \frac{SI - BE}{SI + BE} 100\% \,. \tag{3}$$

The second measure describing the budget balance criterion is the value of the difference between total sellers' income and total buyers' expense related to the market surplus. This measure has got the term budget imbalance in relation to market surplus ($RBIS$):

$$RBIS = \frac{SI - BE}{Q} 100\% \,. \tag{4}$$

Pareto-efficiency. The results given by the trading model are Pareto-efficient, if one can not improve the result for one market participant without making some other participants worse off: the results are Pareto-efficient if such results are not Pareto-dominated by other results.

4.2. Individual Criteria

Individual economic benefits. From the perspective of an individual market participant the trading model should allow to obtain the highest possible value of the individual economic benefits. The measure is defined as the value of the individual utility function for each market participant.

Absolute individual fairness – individual rationality. A trade model is individually rational, if no market participant loses from the participation in the trade. This property is also called the voluntary participation (if it may lose then it can choose not to participate in the trade).

Relative individual fairness. A trade model is fair in a relative sense, when from the perspective of each participant no other offer is favored in relation to its offer.
Other criteria related to fairness have been outlined below:

- anonymity: a participant remains anonymous if renumbering of participants does not affect the obtained outcome;
- symmetry: two participants with the same parameters, in the same market situation (the same utility functions) should obtain the same individual results;
- an equal price: each participant receives the same volume of a commodity for the same unit price.

4.3. Technical Efficiency

The possibility of a practical technical implementation of the trade model is evaluated with technical efficiency indicators dependent on the computational complexity of the model and on its reliability. Exemplary measures are:

– duration of single market balancing;

– total duration of the market balancing (for iterative mechanisms);

– a number of exchanged messages (total, average).

5. Example

The evaluation framework has been applied to test the balancing communication bandwidth trade (BCBT) model discussed in [5]. It is a multicommodity trade model for the market of network transport resources with many buyers and many sellers. The model assumes complete divisibility of offered bandwidth (sell offers) and the ability to allocate any fraction of bandwidth to any bandwidth demand (buy offers). The market balancing process consists of setting the paths serving buy offers with a combination of sell offers. The formal format of the model is a linear programming problem representing a double auction with the maximization of market welfare.
The evaluation of the properties of the BCBT model has been carried out according to the framework principles outlined in the paper. Below a brief description of this test case is presented.

5.1. Classification of the Model

Type of commodities. The BCBT model assumes that both links (sell offers) and paths (buy offers) concern unidirectional network capacity. The division model of the offered bandwidth is continuous: the seller and buyer may indicate bandwidth amount of their bids with any positive real number. The bandwidth can be traded in the range from 0 to a maximal admissible volume specified for the offers.

Relations between offers and commodities. All market offers in the model are elementary offers concerning elementary point-to-point bandwidth connections in the network.

Market participants. The model represents a double auction: there are many buyers and many sellers. It is a centralized trading model – a single market operator balances the market.

Quality of service constraints. There are not any QoS parameters included in the model (e.g., there is no limit on the maximum length of paths).

Resource allocation rule. The balancing mechanism can freely allocate bandwidth in the network, in particular a single sell offer can be used to serve multiple buy offers, a single buy offer can be realized with many paths. There are no restrictions on the possibility of setting paths in the network: there are no predefined paths, they are set during the market balancing process taking into account of all available offers.

Pricing rule. The pricing rule adopted is the dual pricing rule – it is based on dual prices of the balance constraints. In the BCBT model the balance constraints exist only for the links – so the prices are determined directly for the sell offers only. The prices for demands (buy offers) are determined as a sum of the values of bandwidth bought on links in the related paths.

Exchanged messages. The messages exchanged in the model are point characteristics of market participants' preferences in the form of a traditional offer expressing the amount of commodities and their unit prices.

Market balancing dynamics. The model is an example of a one-time double auction – the market participants submit offers and then the market is balanced – resources are allocated and prices are calculated.

Implementation. The model has been implemented as a linear program maximizing total market welfare. The linear program is written in GMPL and AMPL.

5.2. Input Data

The network topology is derived from the example of the NOBEL-EU network from the SNDlib developed in the IST NOBEL project "*Next Generation Optical Networks for Broadband European Leadership*" [11]. The nodes of the network are located in 28 major European cities. The network topology (nodes and potential links) is illustrated in Fig. 3.

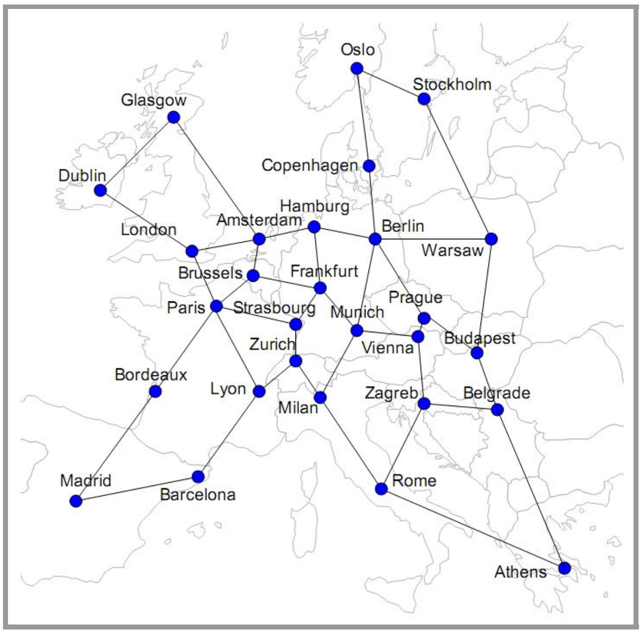


Fig. 3. Pan-European network of EU project IST NOBEL [11].

Bandwidth buy offers are independent from the traffic demands included in the original SNDlib example. It is assumed that the offered prices rise along with the geographical distance between a pair of nodes, and offered bandwidth capacity decreases with the distance. 150 buy offers have been generated by the following algorithm:

1. Draw randomly a source and a destination network node (different from each other).
2. Set pseudo-randomly offer price according to the expression (5) with dependency on the geographical distance between the nodes:

$$\mathrm{Gauss}\left(\mathrm{Avg}P \cdot Dist,\ \frac{\mathrm{Avg}P}{3Dist}\right). \qquad (5)$$

3. Set pseudo-randomly offer bandwidth capacity according to the expression (6) with dependency on the geographical distance between the nodes:

$$\mathrm{Uniform}\big(0.5\,\mathrm{Avg}C,\ 1.5\,\mathrm{Avg}C\big)\left(1-\frac{Dist}{Dist_{\max}}\right) \qquad (6)$$

with the following denotation:

– $\mathrm{Gauss}\,(a,b)$: Gaussian distribution with the expected value a and the standard deviation b;
– $\mathrm{Uniform}\,(a,b)$: uniform distribution in the range $[a,b]$;
– Avg$P$: average unit bandwidth price related to a unit distance;
– Avg$C$: average bandwidth capacity for adjacent nodes (distance equals 0);
– $Dist$: distance between nodes;
– $Dist_{\max}$: the maximal distance between the nodes in the network.

Following parameter values has been used for buy offer generation: Avg$P = 1$, Avg$C = 15$.
Sell offers have been also generated randomly (they have replaced links in the original SNDlib example: the primary links have been interpreted as bandwidth segments on which sell offers are submitted). The assumption made is similar to the case of buy offers: the offered prices rise along with the geographical distance between a pair of nodes, and offered bandwidth capacity decreases with the distance. 5 separate sell offers have been generated on every bandwidth segment using the same mathematical expression for setting prices and capacities as for the buy offers (with the following parameter values: Avg$P = 0.5$, Avg$C = 10$) according to the algorithm:

1. Set pseudo-randomly the direction of the offer (the probability of each direction equals 0.5).
2. Set pseudo-randomly offer price according to the expression (5) with dependency on the geographical distance between the nodes.
3. Set pseudo-randomly offer bandwidth capacity according to the expression (6) with dependency on the geographical distance between the nodes.

The test data generated for the Nobel-EU network including 150 buy offers and 205 sell offers has been saved in a BM-XML format file. The data sat is a simple but complete test case, and it can be used also for testing the properties of other models than the BCBT.

5.3. Adaptation and Running Tests

A single test procedure consists of three phases. The first one is the adaptation of the data stored in BM-XML format to the M3-XML format. The second phase is the solution of the problem in accordance with a given mathematical model of the BCBT. The third phase is the collection of the results and their analysis. The detailed description of each step is as follows:

- The data in the BM-XML format has been converted to the M^3 information model using several XSLT transformations. The result of this procedure is a set of files in the M3-XML format.
- The obtained data set is then passed to the decisional-computational processor performing the following steps: the data is converted to a GMPL model implementing the BCBT allocation model using an XSLT transformation; the complete GMPL model with numeric data is passed to a linear programming solver (e.g., AMPL, GlpSol) returning results.
- The results are parsed to the M3-XML format.

Calculations have been performed on a PC (Intel Core 2 Duo 2.60 GHz, 2 GB RAM).

5.4. Output Data Analysis

The numerical data and short comments for the framework indicators have been presented below.

Economic efficiency. From a global perspective the BCBT model is economically efficient in the sense that it maximizes the global economic welfere. Thus, for given market offers, no better allocation of bandwidth resources is possible. The market welfare achieved in the test case is 13861.

Incentive compatibility. The obtained value of the allocative inefficiency equals 0.026. It is a very good outcome – it indicates that the possibility for speculation on the market is limited.

Budget balance. The RBIT and RBIS benchmarks are 0, so the model fulfills the requirements for the budget balance property.

Pareto-efficiency. The Pareto-efficiency property holds for the family of BCBT models. Pareto-efficiency is ensured if the aggregation function of the individual objective functions is strictly increasing. The family of BCBT models uses objective function as the sum of individual objective functions, which is strictly increasing with respect to every coordinate, so the whole BCBT family is Pareto-efficient.

Individual economic benefits. The calculated contract prices have been used to determine the individual benefits of the market participants. The list of their values has not been included here because of the limited space of the paper.

Absolute individual fairness – individual rationality. The individual benefits for those market participants, whose offers have been successfully traded, are positive. So the requirements of the individual rationality property are fulfilled.

Relative individual fairness. The requirements for relative fairness of individual market participants are partially met in the test case. In general their fulfilment depends on the implementation of solvers (e.g., the order of identical market offers may have impact on the volume of their realization). The BCBT model itself does not have any additional constraints for complying with these requirements.

Duration of market balancing. The total time of the market balancing has been very short: 0.95 s. There have not been any additional computation for resource pricing because the prices have been derived from the parameters of the linear program solution. The LP formulation of the BCBT model does not comprise integer variables, to it is feasible even for large networks.

Number of exchanged messages. The market participants have submitted 150 buy and 205 sell offers – there are total 355 bid messages (the same number of messages covers announcing the results of the trade).

6. Summary

The paper presents an attempt to develop a methodological approach to the problem of comparing different models for trading capacity in the communication transport networks. Several stages of test preparation and running computations have been distinguished and described. The use of standardized data format is essential: the task of comparison of results becomes easier and the same test cases can be reused for many models. The proposed framework will be used to study and compare the properties of different trade models, such as c-SeBiDA (combinatorial sellers' bid double auction) [12], MIDAS [13], BCBT [5], NSP (network service provider) [14], and other new models developed by the authors in their research. The described framework and its tools can be a part of a broader research platform forming an advanced computing environment for testing the market models of resource allocation in communication networks.

The evaluation framework is in its early stage and still there is a need to refine many of its elements. The most important areas perceived for improvement include:

- The development and extension of economic models to describe the bandwidth demand and supply enabling generation of test data reflecting the real market conditions.
- Refine the set of criteria enabling to study and compare specific characteristics of trading models, e.g., resulting from the market game between trade participants.

These and other extensions of the basic methodology proposed in this paper will be presented in the authors' future papers.

Acknowledgements

The authors acknowledge the Ministry of Science and Higher Education of Poland for partially supporting the research through project PBZ-MNiSW-02/II/2007. The results of the research have been discussed within the JRA3.3 activity of the FP7 Euro-NF Network of Excellence.

References

[1] S. S. Chiu and J. P. Crametz, "Taking advantage of the emerging bandwidth exchange market", White paper, RateXchange Ltd., 1999.

[2] W. Stańczuk, J. Lubacz, and E. Toczyłowski, "Trading transport resources of communication networks on bandwidth exchanges", in *Proc. 46th FITCE Congr.*, Warsaw, Poland, 2007.

[3] "SNDlib: survivable fixed telecommunication network design" [Online]. Available: http://sndlib.zib.de/, 2010

[4] "M^3 – multicommodity market model" [Online]. Available: http://www.openm3.org/, 2010

[5] W. Stańczuk, J. Lubacz, and E. Toczyłowski, "Trading links and paths on a communication bandwidth market", *J. Univers. Comput. Sci.*, vol. 14, no. 5, pp. 642–653, 2008.

[6] "Bandwidth pricing report", Tech. Rep., TeleGeography, PriMetrica Inc., 2007.

[7] "XSL transformations" [Online]. Available: http://www.w3.org/TR/xslt, 2010

[8] "Xquery 1.0: an XML query language" [Online]. Available: http://www.w3.org/TR/xquery/, 2010

[9] P. Kacprzak, M. Kaleta, P. Pałka, K. Smolira, E. Toczyłowski, and T. Traczyk, *Procesor decyzyjno-obliczeniowy dla rynkowego modelu danych M^3*. Warsaw: WKŁ, 2008 (in Polish).

[10] "GLPK (GNU linear programming kit)" [Online]. Available: http://www.gnu.org/software/glpk, 2010

[11] "Report on multilayer traffic engineering mechanisms for NOBEL solutions in medium and long term scenarios, deliverable – d2.2", Tech. Rep., IST NOBEL "Next Generation Optical Networks for Broadband European Leadership", Phase 2, 2007.

[12] R. Jain and P. P. Varaiya, "An effcient incentive-compatible combinatorial market mechanism", in *Proc. Allerton Conf.*, Chicago, USA, 2004.

[13] M. Dramitinos, G. D. Stamoulis, and C. Courcoubetis, "An auction mechanism for allocating the bandwidth of networks to their users", *Comput. Netw.*, vol. 51, iss. 18, pp. 4979–4996, 2007.

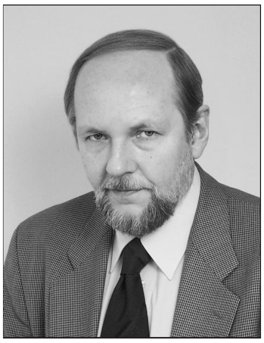
[14] R. Jain and J. Walrand, "An efficient mechanism for network bandwidth auction", in *IEEE Netw. Oper. Manage. Symp. Worksh. 2008*, Salvador da Bahia, Brazil, 208, pp. 227–234.

Wojciech Stańczuk received the M.Sc. degree in telecommunications from Warsaw University of Technology (WUT), Poland, in 2001. Currently he is a Ph.D. student and a researcher in the Institute of Telecommunications (WUT). His research interests lie in modeling techno-economic aspects of communications networks: resource allocation, pricing, infrastructure investments.

e-mail: w.stanczuk@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

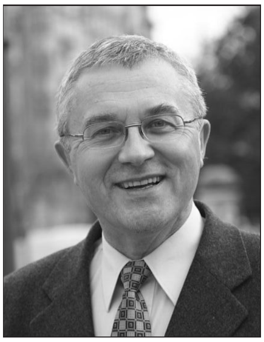
Piotr Pałka is an Assistant Professor of the Operations and Systems Research Division in the Institute of Control and Computation Engineering at the Warsaw University of Technology (WUT), Poland. He received the M.Sc. and Ph.D. degrees in computer science in 2005 and 2009, respectively, both from WUT. His research interest focus on the market mechanisms, especially the incentive compatibility on the infrastructure markets and on the multi-agent systems. His current research is focused on application of multicommodity turnover models.

e-mail: P.Palka@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland



Józef Lubacz received his Ph.D., and D.Sc. at Warsaw University of Technology (WUT), Poland. He was the Head of the Department of Teleinformatics (WUT, 1987–2002), Dean of the Faculty of Electronics and Information Technology (WUT, 2002–2005). Currently he works as a Full Professor at Warsaw University of Technology (WUT). He is a member of the General Council for Higher Education of Poland (since 2006) and the Director of the Institute of Telecommunications (WUT, 2008). His scientific interests focus on: networks planning methodology, non-technological issues of information society and philosophy of technology.

e-mail: jl@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland



Eugeniusz Toczyłowski is a Professor, the Head of Operations and Systems Research Division, at the Institute of Control and Computation Engineering at the Warsaw University of Technology (WUT), Poland. He received the M.Sc. degree in 1973, Ph.D. in 1976, D.Sc. in 1989, and the title of Full Professor in 2004. His main research interests are centered around the operations research models and methods, including structural approaches to large scale and discrete optimization, auction theory and competitive market design under constraints, multicommodity trading models, and design of management information systems.

e-mail: E.Toczylowski@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

Pricing Rules Comparison in the Context of Bandwidth Trade

Piotr Pałka and Eugeniusz Toczyłowski

Abstract—In this paper we compare two pricing rules in the context of bandwidth trade. Allocation and pricing rules, together with a set of signals received from independent agents, constitute a market mechanism. In the paper we analyze two pricing rules: well known Vickrey-Clarke-Groves rule (VCG) and the parametric pricing rule (PPR). We apply these pricing rules to the allocation rule specified by the balancing communication bandwidth trade model (BCBT).

Keywords—communication bandwidth trade, mechanism design, pricing rule.

1. Introduction

Bandwidth market, in the context of independent traders (in this paper we will call them agents), determines a game among these agents. The finite group of agents interacts. The set of agents is denoted $\mathscr{I} = \{1, 2, \ldots, n\}$ and generic agent is represented as i. Every agent tries to maximize her individual profits. Thus, the aims of particular agents are inconsistent. However, the market designer can influence the behavior of particular agents by applying specific rules to the game. Thus, the market designer tries to achieve the overriding goals. In the context of market games, such rules are allocation and pricing rules. Applying this specific rules to the game among independent agents results in the market mechanism.

Every agent is characterized by her preferences. We call this preferences the agent's type T_i. We usually assume that the agent knows only her type, but not those of other agents. The mechanism designer also does not know agents' types. We call such mechanism informationally decentralized [1].

Agent reports the signal θ_i to the mechanism. The signal, in the context of market mechanism, can be understood as a buying/selling offer. Signal is reported on the basis of the strategy function. The strategy function depends on the agent's type, behavior of the other agents and on the mechanism rules. Thus, the signal reported by the particular agent, can differs from this agent's type.

The allocation rule determines the allocation of the offers. It divides the offers for the accepted (also partially accepted) and rejected. Pricing rule sets a vector of prices for traded commodities. The mechanism receives the signals from particular agents and performs the allocation of commodities and determines their valuation – in accordance with the allocation and pricing rules.

2. Desired Mechanisms Properties

Mechanism is constructed in order to fulfill the desired results. Particular agents try to achieve their maximal individual profits, thus their goals are inconsistent. Moreover, goals of particular agents are also inconsistent with global goals, desired by the mechanism designer. Mechanism theory considers the set of the most desired properties.The most important mechanism properties are: incentive compatibility, individual rationality, Pareto-efficiency and budget balance [2].

Incentive compatibility property holds if no agent has incentives to report signal different from her type. In the other words, the incentive compatibility property holds, if no agent has incentives to report untruthful offer. Mechanism is individually rational, if no agent loses from participation in such mechanism. Such property is also called the voluntary participation property. The voluntary participation property means that if agent loses she can choose not to participate in given mechanism. Results of the mechanism are Pareto-efficient, if we can not improve results (i.e., payment) for one agent without making results for some other agents worse off. In the other words, mechanism results are Pareto-efficient if such a results are not Pareto-dominated by other results. Mechanism has balanced budget if sum of sellers expenses is equal to sum of buyers incomes. In the other words, mechanism is budget balanced, if there is no need to surcharge mechanism, and the mechanism does not give us additional profits.

The most desired mechanism should fulfill all listed properties. However, several theorems [3]–[5] (so called impossibility theorems) states, that it is impossible that single mechanism holds all properties. Therefore, we need to decide for the subset of desired mechanism properties, which should hold for considered mechanisms.

In the following parts of the paper, we introduce and compare two mechanisms for the bandwidth trade problems. These mechanisms are Pareto-efficient, individually rational and incentive compatible. However, we allow for the lack of budget balance.

3. Mechanisms Review

Many papers deal with the bandwidth trade problems. We will describe a number of them using the mechanism theory terminology.

In this section we review the variety of bandwidth trading mechanisms presented in the literature. There are three classes of such approaches:

– simultaneous, single link auctions [6], [7] – here we discuss the most recent mechanism MIDAS [6];

– combinatorial auctions [8], [9], with the c-SeBiDA combinatorial double auction as a good representative;

– the family of multicommodity market models (M^3), with basic BCBT market model [10].

3.1. Simultaneous, Single Link Auctions

In the simultaneous, separate auctions for individual links an agent that wants to buy a certain path must put simultaneous bids at all relevant auctions. Then special, iterative mechanisms are required to coordinate individual links. This aspect, as well as possible suboptimality are the main roots of our criticisms for these methods. The review of the auction mechanisms dealing with the problem of coordination of simultaneous, single link auctions is presented in [6]. Authors point out several drawbacks of already proposed mechanism, such as the convergence problem and lack of incentive for submitting truthful bids. In the [6] the simultaneous multi-unit Dutch auction one for each link mechanism (MIDAS) is proposed. The allocation rule of the MIDAS derives from the generalized Vickrey auction [11], but is carried out by the simultaneous Dutch auctions. The payment rule of the mechanism is equivalent to Vickrey-Clarke-Groves payment rule. However, we have to remember that incentive compatibility property is satisfied only when allocation rule is efficient, which is not always true in the case of MIDAS mechanism. Thus, even thought the mechanism may seem to be simple and scalable, the complicated synchronization, that requires full information makes it impractical in our opinion.

3.2. Combinatorial Auctions

Combinatorial auctions are designed for trading on dependent commodities. One particular auction model that appears in the context of bandwidth market is the combinatorial seller's bid double auctions (c-SeBiDA) [8]. The c-SeBiDA considers two types of commodities: inter-node links and paths consisting of particular links. Agents may bid a single link or a bundle of links constituting specific path. Allocation rule ensures that the same indivisible amount of bandwidth is assigned to all links constituting buyers path, thus a buyer has no risk of buying different amount of bandwidth on some required links. The c-SeBiDA auction has several valuable properties, such as the maximization of the global economic wealth. However, similarly to the approaches concerning on simultaneous auctions, buyers bids must specify the particular links that constitute a desired path. This may lead to welfare inefficiency. Welfare inefficiency corresponds to a situation where the social welfare obtained is not maximum possible to reach – we can imagine the allocation rule, which allocates resources in a better way.

The c-SeBiDA mechanism is individually rational. Its results are Pareto-efficient and it has balanced budget. However, such a mechanism does not hold incentive compatibility property – particular agents can derive unreasonable profits from this mechanism.

3.3. Multicommodity Market Models

In the multicommodity auction models, the efficient market balance is obtained in the effect of joint optimization of many elementary buy and sell offers. Multicommodity means that market entities can trade with bundles (packages) of different commodities. The balancing communication bandwidth trade (BCBT) model proposed in [10] allows bidders to place buy offers not only for bundled links, but rather for end-to-end connections. Therefore, no buyer does not have to know which links to choose to best allocate the demanded capacity. It is the decision model that allocates the most efficient links to paths.

We assume that the communication network consists of nodes connected by links. The inter-node link represents a network resource (bandwidth), that can be an elementary commodity offered for sale on the bandwidth market. However, network resources being traded can be more complex and can be composed of many parallel links, or end-to-end node connections represented by paths or subnetworks.

Each buy offer concerns a point-to-point bandwidth connection between a pair of specified locations in a communication network. The locations form the set of network nodes $\mathscr{V}$. The connections (and links) are unidirectional, i.e., they have source and sink nodes.

The objective of BCBT model is the maximization of total economic welfare Eq. (1), which is the total surplus of all buyers and sellers. Constraints (2) and (3) set upper and lower bounds on particular network links (x_e) and particular end-to-end network demands (x_d). The non-negative variable x_{ed} constraint (4) is interpreted as a bandwidth capacity allocated to network link e to serve end-to-end demand d. Also, the sum of capacities allocated to all network demands $\sum_{d\in D} x_{ed}$ served by particular network link e, should not exceed the realization x_e of the link constraint (5). Finally, the sum of all capacities, provided with incidence matrix a_{ve}, allocated to all network links, serving particular network demand, should not exceed the realization of the end-to-end demand x_d Eq. (6):

$$\hat{Q} = \max\left(\sum_{d\in\mathscr{D}} E_d x_d - \sum_{e\in\mathscr{E}} S_e x_e\right), \tag{1}$$

subject to:

$$0 \le x_d \le h_d, \quad \forall d \in \mathscr{D}, \tag{2}$$

$$0 \le x_e \le y_e, \quad \forall e \in \mathscr{E}, \tag{3}$$

$$\sum_{d\in\mathscr{D}} x_{ed} \le x_e, \quad \forall e \in \mathscr{E}, \tag{4}$$

$$0 \le x_{ed}, \qquad \forall e \in \mathscr{E}, \quad \forall d \in \mathscr{D}, \tag{5}$$

$$\sum_{e\in\mathscr{E}} a_{ve}x_{ed} = \begin{cases} x_d & v = s_d \\ 0 & v \neq s_d, t_d \\ -x_d & v = t_d \end{cases}, \forall v \in \mathscr{V}, \forall d \in \mathscr{D}, \quad (6)$$

where:

indices:

$d \in \mathscr{D}$ buy offers – demands for bandwidth,
$v \in \mathscr{V}$ network nodes,
$e \in \mathscr{E}$ sell offers – network resources;

parameters:

a_{ve} $= 1$ if link e originates in node v,
$= -1$ if e terminates in node v,
$= 0$ otherwise,
s_d source node for demand d,
t_d sink node for demand d,
h_d required capacity of demand d,
E_d offered unit price for demand d,
y_e offered capacity of network link e,
S_e offered unit price for network link e;

variables:

x_{ed} bandwidth flow serving demand d allocated to network link e,
x_d contracted bandwidth capacity for demand d,
x_e contracted bandwidth capacity for network link e.

The x_e and x_d are, respectively, values of realized bandwidth on the link e and on the demand d. They are also the accepted offers for link e and demand d – in the BCBT model sell offers correspond network links and buy offers correspond demand paths resulting in a multigraph.

As stated before, we will identify offerers with agents, and the submitted offers with signals send to mechanism. Because single offer relates to single link or single path, we can identify the offers with the network resources. Thus, let us define the set of agents as the sum of the sets $\mathscr{I} = \mathscr{E} \cup \mathscr{D}$. Also let us define the signal θ_i sent by the ith agent as the tuple: offered price and offered capacity. When ith agent represents bandwidth seller, such tuple is equal to $\theta_i =< S_i, y_i >$, otherwise such tuple is equal to $\theta_i =< E_i, h_i >$. Also the allocation results we will denote as $\boldsymbol{x} = (x_1, \cdots, x_i, \cdots, x_n)$.

Let us note that the BCBT model does not define any pricing rule, it defines only allocation rule. So we will treat the BCBT model as the allocation rule. This also applies to other bandwidth trade models from the BCBT family. Therefore, we need to propose the pricing rules to the base BCBT allocation rule. As the result we obtain two market mechanisms presented in the next sections.

4. Analyzed Pricing Rules

As we stated before, we analyze and compare two pricing rules in the context of bandwidth trade. These rules are the Vickrey-Clarke-Groves (VCG) rule (mechanism) and the parametric pricing rule (PPR).

4.1. Vickrey-Clarke-Groves Pricing Rule

VCG pricing rule – or rather VCG mechanism was introduced in the papers [12]–[14]. VCG mechanism does not define allocation rule, it only states that applying VCG pricing rule to efficient allocation rule creates VCG mechanism. An allocation rule is said to be efficient if it maximizes "social welfare", treated as the aggregation of particular agents' utility functions. Thus, the BCBT is an efficient allocation rule.

For the sake of simplicity, let us assume, that every agent submits one offer. VCG pricing rule sets payoff for every agent. Payoff for ith agent is defined as the opportunity cost that the presence of ith agent introduces to all other agents.

Set of agents sends a vector of signals $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$ to the mechanism. Let us define vector $\boldsymbol{\theta_{-i}} = (\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_n)$, which contains the set of all signals but signal θ_i. The $Q(\boldsymbol{\theta})$ is the economic welfare obtained by the allocation rule with all the signals $\boldsymbol{\theta}$. The $Q(\boldsymbol{\theta_{-i}})$ is the economic welfare, obtained by the allocation rule without ith signal. So, the payment for ith agent is equal to:

$$I_i = Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta_{-i}}) \quad (7)$$

We can determine prices for each agent (if her offer is accepted). If the agent i submitted the selling offer, the price for her is equal to Eq. (8), otherwise, the buying price for ith agent is equal to Eq. (9). Prices for agents which offers were rejected are negligible.

$$\pi_i^S = \frac{S_i x_i + I_i}{x_i}, \quad (8)$$

$$\pi_i^K = \frac{E_i x_i - I_i}{x_i}. \quad (9)$$

Applying the VCG pricing rule to the BCBT allocation rule results in the VCG mechanism. Properties of such mechanism are as follows: it fulfills incentive compatibility property, results of the mechanism are Pareto-efficient, individual rationality is also fulfilled. Unfortunately, there is a lack of budget balance.

4.2. Parametric Pricing Rule

Second considered pricing rule is the parametric pricing rule [15]. This rule sets prices accordingly to modified Vickrey double auction (MVDA) [16]. The MVDA mechanism was designed for double auction with indivisible commodities. It sets the differentiated buying and selling prices in such a way, that there is no agent, which has incentives to deviate from her type.

Parametric pricing rule uses parametric analysis performed in the respect to the mathematical model of the allocation rule. Parametric analysis is based on repeatedly performed sensitivity analysis. Sensitivity analysis also is performed in the respect to the mathematical model of the allocation rule. Sensitivity analysis provides us with results, which tell us how much the ith offer price can be changed without changing the commodities allocation. Let us denote results performed in step κ of parametric analysis as $s^{(\kappa),+}$ for the selling prices, and as $e^{(\kappa),-}$ for the buying prices.

Given ith price (for the sake of simplicity let us assume that it is selling price) is increased of the value $s_i^{(\kappa),+} + \varepsilon$ (where $0 < \varepsilon \ll 1$). Afterwards, the allocation model is solved again and the allocation of the commodities changes. Particular steps of parametric analysis set more beneficial price for ith offer, nevertheless accepted volume of ith offer decreases. The analysis is performed until given ith offer is rejected. On the basis of ith offer price in the last step (let us denote the number of the last step by κ^*) the individual price (π_i^S or π_i^K) for such offer is set.

Parametric pricing rule sets individual prices for each offer. Combined with the BCBT allocation rule constitutes a mechanism. This mechanism fulfills incentive compatibility property, its results are Pareto-efficient, individual rationality is also fulfilled.

Selling prices set by the parametric pricing rule are not lesser than buying prices. Thus the budget balance property does not hold for mechanisms with the parametric pricing rule.

4.3. Imbalance Reduction

We propose the algorithm to reduce the budget imbalance. The main idea of this algorithm is to change the stop criterion of the parametric pricing rule. Stop criterion of the parametric rule generally implies that the parametric analysis will be carried out until the rejection of given offer. We propose change of the stop criterion – the analysis will be carried out until given offer is profitable:

$$I_i = x_i(\pi_i^S - S_i^0) \qquad \forall_{i \in E}, \tag{10}$$

$$I_i = x_i(E_i^0 - \pi_i^K) \qquad \forall_{i \in D}. \tag{11}$$

However, to perform such algorithm modification, we have to calculate profit for every agent. Equations (10) and (11) represent the profit of agents. Let us notice, that the rules to calculate the profits, need to know agents' types. However, in the previous sections we have assumed that types are private knowledge of particular agents. Nevertheless, we can assume that required value of S_i^0 or E_i^0 (prices that correspond to the agents' types) belongs to certain intervals. Such intervals can be determined on the basis of expert and common knowledge, historical data, etc. Therefore, let us assume that these values belong to the following intervals $S_i^0 \in <\underline{S}_i^0, \overline{S}_i^0>$ and $E_i^0 \in <\underline{E}_i^0, \overline{E}_i^0>$.

The main idea of the algorithm for improve the budget balance is to limit price for particular participants. Limitation of price shall be made by reducing the number of steps of the parametric analysis. The algorithm retains good mechanism properties: incentive compatibility, individual rationality and Pareto-efficiency, and reduces the budget imbalance.

5. Experimental Studies

First, on the simple case study, we show differences and similarities between the VCG and the PPR pricing rules. Next, we present a series of experiments. Such experiments focus on three networks from the SNDlib [17] repository. On such experiments we compare the imbalance measure.

5.1. Simple Case Study

The simple case study was performed on the exemplary four-node network (see Fig. 1).

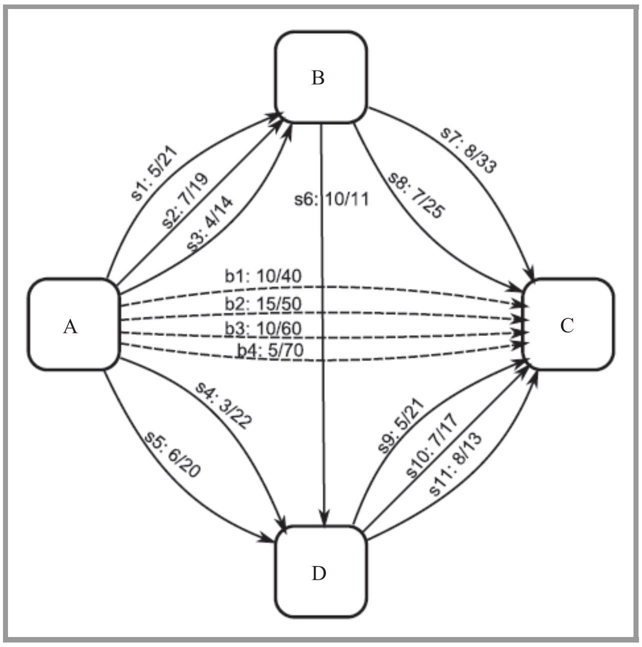


Fig. 1. Four-node network. Solid lines represent links, dotted lines represent demands. The notation is following: *o:v/p*, where *o* means offer ID, *v* means offered bandwidth volume and *p* means offer price for unit of the bandwidth.

In Table 1 we can see prices, set by the mechanisms applying parametric rule and the VCG pricing rule, to particular agents. We can also see results of allocation rule. Some of prices are the same for VCG and PPR, some prices are greater for VCG, some are greater for PPR, nevertheless all the prices are very similar.

Both mechanisms (applying the parametric pricing rule and applying the VCG pricing rule) do not meet the budget balance property. To compare such mechanisms, we propose the measure – relative budget imbalance ratio Eq. (12). Proposed measure reflects the degree of non-compliance with budget balance property. The SI is the total seller's incomes $SI = \sum_{i \in E} S_i x_i$, the BE is the total buyer's expenses $BE = \sum_{i \in D} E_i x_i$:

$$RBUT = \frac{SI - BE}{SI + BE}. \tag{12}$$

Relative budget imbalance ratio for the mechanism applying the parametric pricing rule is equal $RBUT = 3.79\%$, and

Table 1
Prices comparison for mechanisms applying the parametric pricing rule and VCG pricing rule

Nodes	Offer e	Offered vol. S_e	Acc. vol. y_e	Offer price x_e	VCG π_e^S	Comp.	PPR π_e^S
	–	[Mbit/s]		[Euro/Mbit/s]			
A-B	s1	5	2	21	22	=	22
	s2	7	7	19	21.57	>	21
	s3	4	4	14	21.25	>	21
A-D	s4	6	6	20	32.50	>	32
	s5	3	3	22	32	=	32
B-D	s6	10	6	11	13.33	<	14
B-C	s7	7	7	29	30.54	<	31
	s8	8	0	32	–		–
D-C	s9	7	7	17	19.57	<	20
	s10	8	8	13	19.75	<	20
	s11	5	0	21	–		-
	d	E_d	h_d	x_d	π_d^K		π_d^K
	–	[Mbit/s]		[Euro/Mbit/s]			
A-C	b1	10	10	60	49.80	<	50
	b2	15	7	50	47.17	>	47
	b3	5	5	70	50	=	50
	b4	10	0	40	–		–

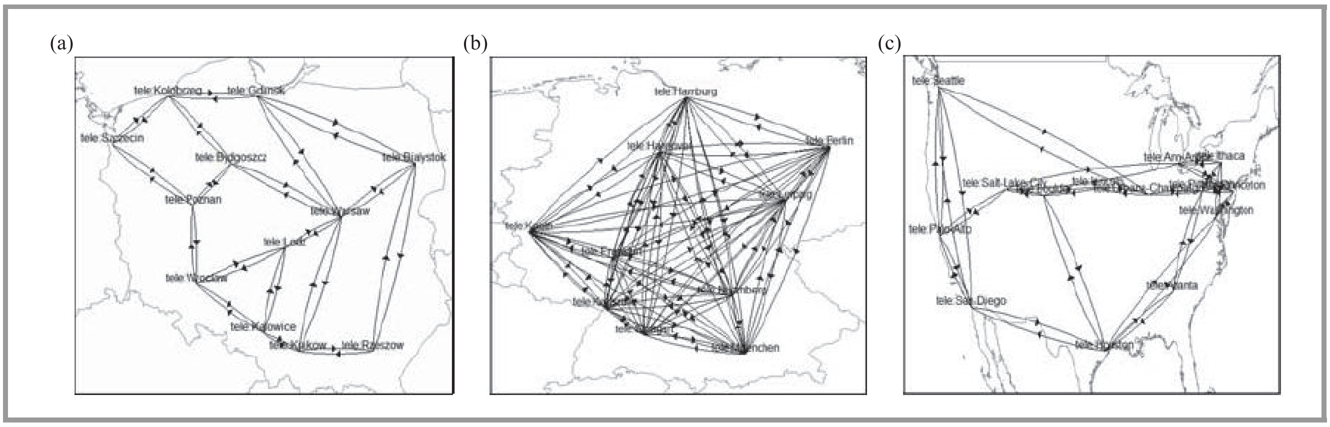


Fig. 2. Network instances [17]: (a) Polska; (b) dfn-bwin; (c) nobel-us.

for the mechanism applying the VCG pricing rule is equal to $RBUT = 3.64\%$ (see Table 1). We can see, that the measure is slightly better for mechanism applying the VCG pricing rule.

5.2. Series of Experiments

Three series of experiments were performed, each contained ten experiments. The experiments concern bandwidth trading performed on the networks taken from the SNDlib library [17]. We analyzed following network instances: Polska, dfn-bwin and nobel-us (Fig. 2). We generated offers for particular network resources (links and demands). Offer prices for particular links and demands was generated on the basis of absolute value of normal distribution, where mean of the distribution was equal to distance between end-to-end nodes. The maximal accepted

Table 2
Aggregated relative budget imbalance for the series of experiments performed on the three backbone networks

Agg. measure		VCG	PPR
Polska			
RBUT	min	0.28	0.37
	mean	0.33	0.45
	max	0.44	0.53
dfn-bwin			
RBUT	min	0.35	0.59
	mean	0.43	0.65
	max	0.47	0.73
nobel-us			
RBUT	min	0.34	0.46
	mean	0.38	0.52
	max	0.40	0.61

trade volumes were generated from absolute value of normal distribution.

In Table 2 we can see aggregated results of the experiments. As we can see, for analyzed cases, the mean value of the relative budget imbalance is lesser for the VCG pricing rule. Thus, we can state, that the VCG pricing rule is better than the parametric pricing rule in terms of the imbalance measure.

6. Summary

In the paper we have analyzed two pricing rules in the context of the balancing communication bandwidth trade allocation rule. These pricing rules are the VCG pricing rule and the parametric pricing rule. Both rules have good properties for multicommodity exchange with infrastructure constraints, specifically for the bandwidth trade. Both pricing rules have cost for adopting, which results from budget imbalance. The experiments show, that for the given data, budget imbalance measure is better for the VCG pricing rule.

Acknowledgments

The authors acknowledge the Ministry of Science and Higher Education of Poland for partially supporting the research through project PBZ-MNiSW-02/II/2007.

References

[1] L. Hurwicz and S. Reiter, *Designing Economic Mechanisms*. New York: Cambridge University Press, 2006.

[2] V. Krishna, *Auction Theory*. San Diego: Academic Press, 2002.

[3] L. Hurwicz, "On informationally decentralized systems", in *Decision and Organisation: A Volume in Honour of Jacob Marchak*. 1972, pp. 297–336.

[4] B. Holmström, "Moral hazard in teams", *Bell J. Econom.*, vol. 13, no. 2, pp. 324–340, 1982.

[5] R. Myerson and M. Satterthwaite, "Efficient mechanisms for bilateral trading", *J. Econom. Theory*, vol. 28, pp. 265–281, 1983.

[6] M. Dramitinos, G. D. Stamoulis, and C. Courcoubetis, "An auction mechanism for allocating the bandwidth of networks to their users", *Comput. Netw.*, vol. 51, no. 18, pp. 4979–4996, 2007.

[7] A. Lazar and N. Semret, "Design and analysis of the progressive second price auction for network bandwidth sharing", Tech. Rep., Columbia University Center for Telecommunications Research, 1999.

[8] R. Jain and P. Varaiya, "An efficient incentive-compatible combinatorial market mechanism", in *Proc. Allerton Conf.*, Chicago, USA, 2004.

[9] R. Jain and J. Walrand, "An efficient mechanism for network bandwidth auction", in *Proc. Netw. Oper. Manage. Symp. NOMS 2008 Worksh.*, Salvador da Bahia, Brazil, 2008, pp. 227–234.

[10] W. Stańczuk, J. Lubacz, and E. Toczyłowski, "Trading links and paths on a communication bandwidth market", *J. Univers. Comput. Sci.*, vol. 14, no. 5, pp. 642–653, 2008.

[11] L. Ausubel, A Generalized Vickrey Auction. 2000.

[12] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders", *J. Finan.*, vol. 16, no. 1, pp. 8–37, 1961.

[13] E. H. Clark, "Multi-part pricing of public goods", *Publ. Choi.*, vol. 11, no. 1, pp. 17–33, 1971.

[14] T. Groves, "Incentives in teams", *Econometrica*, vol. 41, no. 4, pp. 617–631, 1973.

[15] P. Pałka and E. Toczyłowski, "Mechanizmy wyceny dóbr za pomocą uogólnionej metody Yoona i metody analizy parametrycznej", *Automatyka*, vol. 13, no. 2, pp. 539–550, 2009 (in Polish).

[16] K. Yoon, "The modified Vickrey double auction", *J. Econom. Theory*, vol. 101, no. 2, pp. 572–584, 2001.

[17] SNDlib: Survivable fixed telecommunication network design" [Online]. Available: http://sndlib.zib.de/

Piotr Pałka, Eugeniusz Toczyłowski – for biography, see this issue, p. 60.

Paper

Bandwidth Trading: A Comparison of the Combinatorial and Multicommodity Approach

Kamil Kołtyś, Piotr Pałka, Eugeniusz Toczyłowski, and Izabela Żółtowska

Abstract—Since the telecommunication market becomes more complex and dynamic, a strong need for a new, efficient and flexible bandwidth trading mechanisms appears. We believe that good mechanisms, that allow effective and fair allocation of bandwidth between market participants will help to develop the real competitive bandwidth market. In this paper we compare two different double-sided bandwidth auction mechanisms, that seem to be well suited approaches for trading indivisible units of bandwidth: combinatorial auction c-SeBiDA and multicommodity mechanism BACBR-I. The c-SeBiDA mechanism considers two types of commodities: inter-node links and paths consisting of particular links. Market participants may bid a single link, or a bundle of links, constituting a specific path. The BACBR-I mechanism is a multicommodity exchange model, that allows bidders to place buy offers not only for individual or bundled links, but rather for end-to-end connections. Therefore, it is the decision model that allocates the most efficient links to connections. We run a large set of experiments to test the allocation and computational efficiency obtained under both approaches.

Keywords—bandwidth allocation, combinatorial auction, computational efficiency, indivisible resources, multicommodity trade.

1. Introduction

We consider a multilateral network resources market. The market is supplied by many participants such as companies laying cables, network providers and other telecommunication link owners. The customers of the market are service providers (ISP, ASP, etc.), geographically spread organization and also network providers who want to expand their network coverage. We assume that sellers offer single telecommunication links and buyers want to purchase bandwidth between two nodes that may not necessarily be directly connected by single link. Requirement of trading end-to-end connections makes the allocation problem combinatorial, because bandwidth demand can be realized by several network links.

After the debacle of Enron Broadband Services in fall 2001 the development of organized market for bandwidth slowed down. Currently, the dominating form of bandwidth trading are bilateral agreements in which two participants negotiate the contract terms. The negotiations are complex, nontransparent and time consuming. This form of bandwidth trading requires a business relationships and often it is inefficient both globally and individually (especially for participants that have not relevant business relationships). The buyer that wants to purchase bandwidth between two nodes connected by a sequence of links owned by different providers must independently negotiate with all of them. If the negotiation fails with one of them (whereas agreements with other sellers would be drawn up and signed), the buyer will get useless bandwidth as it will not ensure the connection between selected nodes. Also even if the buyer manage to purchase bandwidth along some path connecting chosen nodes, there is a risk that this path would not be the cheapest one from all existing paths between this nodes. Thus there is a need of designing more sophisticated market mechanisms that will not have such severe drawbacks that are involved with bilateral agreements. Lately analysis of bandwidth market collapse in 2001 gives promise of emerging new forms of bandwidth trading in the future thanks to especially technologies like global managed private line (GMPLS) and automatic switched optical network/automatic switched transport network (ASON/ASTN) [1], [2].

In this paper we focus on auction based market mechanisms. We analyze two bandwidth auctions: combinatorial sellers' bid double auction (c-SeBiDA) [3] and model for balancing aggregated communication bandwidth resources with indivisible constraint (BACBR-I) that is an extension of balancing aggregated communication bandwidth resources (BACBR) model [4]. Our aim is to compare the allocation and computational efficiency of aforementioned mechanisms. In Section 2 we present considered mechanisms in terms of general properties of auction and the applied approach of supporting end-to-end connection trading. In Section 3 we formulate mathematical models of c-SeBiDA and BACBR-I. Section 4 contains experimental results. Section 5 summarizes our findings.

2. Bandwidth Auction Properties

2.1. General Properties

There are different types of auctions. The c-SeBiDA and BACBR-I mechanisms can be classified according to the auction taxonomy presented in [5] as single-round, socially efficient and double-sided auction of indivisible goods.

Single-round auction are conducted in a single step. Bandwidth market participants submit their offers and the auc-

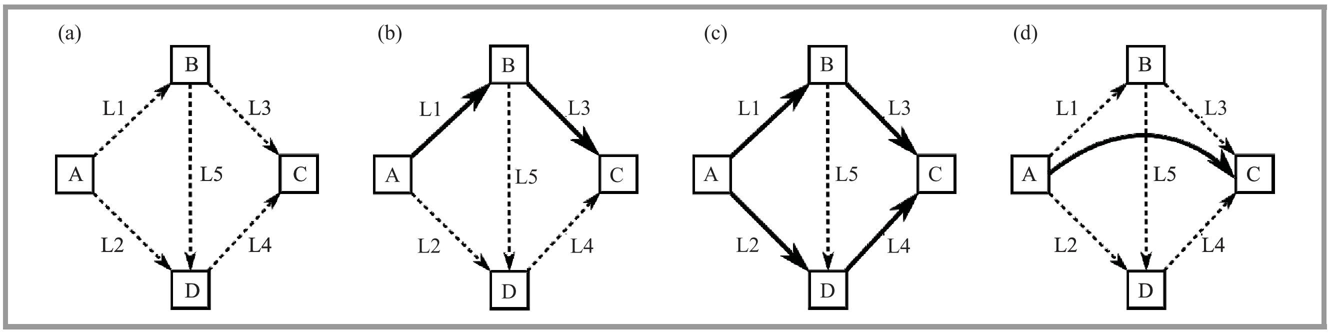


Fig. 1. Different types of supporting end-to-end connection trading: (a) network topology; (b) explicit single path specification; (c) explicit set of admissible paths specification; (d) implicit all possible path specification.

tion mechanism determine an allocation. Other type of auction is progressive one that is carried out in rounds. Progressive auction of bandwidth is proposed in [6].

Double-sided auction concerns multilateral exchange. It can be applied on the bandwidth market where there are many sellers and many buyers. Models for one-sided bandwidth auction (with one seller and many buyers) can be found in [5]–[7].

Auction is socially efficient when it aims at maximizing social welfare. Social efficiency is usually a goal of market mechanisms for bandwidth trading, especially for double-sided auctions. In case of one-sided bandwidth auction also other goals are taken into account, i.e. maximization of seller's revenue [5], [6].

Indivisible goods are integral and cannot be exchanged partially. Bandwidth may be treated as indivisible (modular) commodity. In real networks links often consist of modules that refer to specific transmission and encoding or framing schemes. This modules have determined capacity, i.e., T1 – 1.52 Mbit/s, E1 – 2.04 Mbit/s, OC-3 – 155.52 Mbit/s [8]. Above data transfer standards tend to form standardized contracts on bandwidth market that also use pre-specified amounts of bandwidth [9]. However there are market mechanisms that assume that bandwidth is available in any real fraction of Mbit/s [7], [10].

2.2. Supporting End-to-End Connection Trading

In case of bandwidth auction the essential property is how it supports end-to-end connections trading. We can distinguish three approaches: explicit single path specification, explicit set of admissible paths specification, implicit all possible path specification. All this approaches are illustrated in Fig. 1. Consider network presented in Fig. 1(a) and suppose that buyer wants to purchase bandwidth between nodes A and C. This end-to-end connection has three possible realizations by following sequences of links: L1–L3, L2–L4 and L1–L5–L4.

The first way of supporting end-to-end connection trading relies on explicitly specifying a single path that connects selected nodes. In considered example it can be a path L1–L3 (see Fig. 1(b)). The mechanism should guarantee that the same amount of bandwidth will be allocated at link L1 and L2. This approach is employed by the c-SeBiDA that enables the buyer to submit offer concerning bundle of links constituting particular path. Thus c-SeBiDA is a combinatorial auction. However, explicit single path specification can be also implemented in different manner, i.e., by simultaneous multi-unit dutch auctions [6].

The second approach is more flexible than the first one because it enables to specify a set of admissible paths that can be used to realize end-to-end connection. In considered example buyer may stipulate two paths L1–L3 and L2–L4 (see Fig. 1(c)). The mechanism may allocate different amount of bandwidth at each path, i.e., 90% of demanded bandwidth at path L1–L3 and the rest 10% at path L2–L4, but the summary bandwidth allocated at all paths must be not greater than the buyer's demand. This way of supporting end-to-end connection trading is proposed in [5], [7].

The last approach is the most flexible from the buyer point of view as it enables implicitly specification of all possible paths by submitting offer directly at pair of nodes posing the source and target of end-to-end connection. In considered example, buyer specifies in the offer only source and target of end-to-end connection – appropriately nodes A and C (see Fig. 1(d)). The mechanism itself decides which links are used for realizing this demand. The end-to-end connection between nodes A and C may be realized by several sequence of links, i.e., 40% of demanded bandwidth is served by L1–L3, 50% by L2–L4 and 10% by L1–L5–L4. The BACBR-I applies this approach. It defines two types of commodities: links that are offered for sale and end-to-end connections which are the subject of demand. The buyer that submits an offer for end-to-end connection gets bandwidth at links that generally pose a flow between two selected nodes.

2.3. Auction Rules

Auction mechanism specifies information that market participants must include in their offers. This information is leveraged by two rules of mechanism: allocation rule and pricing rule. Allocation rule decides which offers are accepted. Pricing rule defines the buyers' charges and sellers'

incomes. Both this rules are necessary for clearing the market.
The c-SeBiDA and BACBR-I are single-round, socially efficient and double-sided auction of bandwidth considered as indivisible good. The essential difference between this two mechanisms is in the way of supporting trading end-to-end connections. Our goal is to study how this different approaches affects allocation and computational efficiency of this mechanism. Thus further we analyze only allocation rule of considered mechanisms because it is the one that implements the method of end-to-end connection trading and responds for determining optimal value of the social welfare. Nonetheless, it is worth to mention that pricing rule also affects mechanism efficiency as it decides if mechanism gives incentives for truthful bidding. Here we assume that market participants are truthful, so the allocation rule is maximizing substantial social welfare.

3. Mathematical Models

3.1. Notation

The c-SeBiDA and BACBR-I allocation rules can be formulated as mixed integer linear problems. Below we present notation used in both mathematical programming models:

sets:

- B buy offers,
- S sell offers,
- E network links,
- $S(e)$ sell offers concerning link $e \in E$, $S(e) \subseteq S$,
- D end-to-end connections[1],
- $B(d)$ buy offers concerning end-to-end connection $d \in D$, $B(d) \subseteq B$[1],
- V network nodes[1]

parameters:

- $z_m^{\max}$ maximum units of bandwidth offered for purchase according to buy offer $m \in B$,
- E_m unit price of buy offer $m \in B$,
- $z_l^{\max}$ maximum units of bandwidth offered for sale according to sell offer $l \in S$,
- S_l unit price of sell offer $l \in S$,
- b_{em} $= 1$ if link $e \in E$ belongs to bundle for which buy offer $m \in B$ is submitted, $= 0$ otherwise[2],
- M_d indivisible unit size of end-to-end connection $d \in D$[1],
- M_e indivisible unit size of link $e \in E$[1],
- a_{ve} $= 1$ if node $v \in V$ is source of link $e \in E$, $= -1$ if node $v \in V$ is target of $e \in E$, $= 0$ otherwise[1],
- s_d source of end-to-end connection $d \in D$[1],
- t_d target of end-to-end connection $d \in D$[1];

[1] Only relevant to BACBR-I model.
[2] Only relevant to c-SeBiDA model.

variables:

- z_m realization of buy offer $m \in B$,
- z_l realization of sell offer $l \in S$,
- x_{ed} bandwidth flow serving end-to-end connection $d \in D$ allocated to network link $e \in E$[1].

Both c-SeBiDA and BACBR-I collect all buy (B) and sell (S) offers. They require that each buy offer $m \in B$ contains the maximum buy unit price E_m and the maximum units of bandwidth offered for purchase $z_m^{\max}$. Similarly each sell offer $l \in S$ has to include the minimum sell unit price S_l and the maximum units of bandwidth offered for sale $z_l^{\max}$. According to this two parameters of submitted offers considered mechanisms determine the optimal allocation specified by variables z_m and z_l.

3.2. The c-SeBiDA Model

The c-SeBiDA model assumes that sell and buy offers concern network links (E). Sell offer regards single link. For each link many sell offers can be submitted ($S(e)$). Buy offer is combinatorial and it regards bundle of links (defined by parameters b_{em}).
The mathematical model of the c-SeBiDA is following:

$$\hat{Q} = \max\left(\sum_{m\in B} E_m z_m - \sum_{l\in S} S_l z_l\right), \tag{1}$$

$$0 \le z_m \le z_m^{\max}, \quad \forall_{m\in B}, \tag{2}$$

$$0 \le z_l \le z_l^{\max}, \quad \forall_{l\in S}, \tag{3}$$

$$z_m \in \mathbb{Z}, \quad \forall_{m\in B}, \tag{4}$$

$$\sum_{m\in B} b_{em} z_m \le \sum_{l\in S(e)} z_l, \quad \forall_{e\in E}. \tag{5}$$

The aim of c-SeBiDA is maximizing social welfare. Thus objective function is defined by Eq. (1), where $\hat{Q}$ denotes optimal value of social welfare. First two constraints (2) and (3) set lower and upper bounds of buy and sell offers realizations, respectively. Next constraint (4) ensures that buy offer realization is integral. Sell offers realization will be integral due to constraint (5) which assure that aggregated demand for link is not greater than aggregated supply of bandwidth for that link. More details about c-SeBiDA can be found in [3].

3.3. The BACBR-I Model

The BACBR-I model considers two types of commodities: network links (E) and end-to-end connections (D). Each link and end-to-end connection has predefined module size in which its bandwidth can be traded (respectively, M_e and M_d parameters). Sell offer regards single link and for each link many sell offers can be submitted ($S(e)$). Buy offer concerns single end-to-end connection and for each connection many buy offers can be submitted ($B(d)$). Because

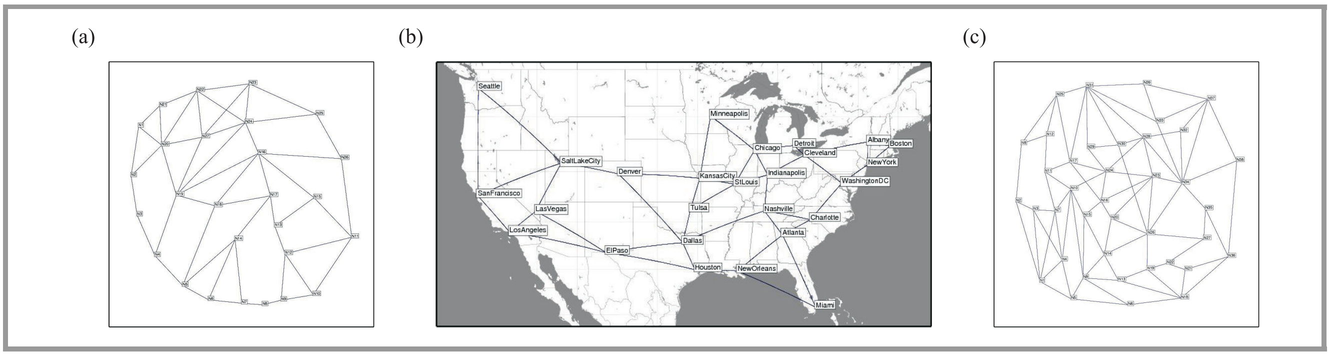


Fig. 2. Network topologies [11]: (a) network sun; (b) network janos-us; (c) network giul39.

BACBR-I itself chooses the links that realize particular end-to-end connection (x_{ed}), it must have an information about network topology. Thus, a set of network nodes (V) and incidence matrix elements (a_{ve}) are given. Also for end-to-end connections source (s_d) and target (t_d) nodes are specified.

The mathematical model of the BACBR-I is following:

$$\hat{Q} = \max\left(\sum_{m\in B} E_m z_m - \sum_{l\in S} S_l z_l\right), \quad (6)$$

$$0 \le z_m \le z_m^{\max}, \quad \forall_{m\in B}, \quad (7)$$

$$0 \le z_l \le z_l^{\max}, \quad \forall_{l\in S}, \quad (8)$$

$$z_m \in \mathbb{Z}, \quad \forall_{m\in B}, \quad (9)$$

$$z_l \in \mathbb{Z}, \quad \forall_{l\in S}, \quad (10)$$

$$\sum_{d\in D} x_{ed} \le \sum_{l\in S(e)} M_e z_l, \quad \forall_{e\in E} \quad (11)$$

$$0 \le x_{ed}, \quad \forall_{e\in E, d\in D} \quad (12)$$

$$\sum_{e\in E} a_{ve} x_{ed} = \begin{cases} \sum_{m\in B(d)} M_d z_m & v = s_d \\ 0 & v \neq s_d, t_d \\ -\sum_{m\in B(d)} M_d z_m & v = t_d \end{cases}, \quad \forall_{v\in V, d\in D} \quad (13)$$

First four equations in BACBR-I model Eqs. (6)–(9) are the same as in the c-SeBiDA. It stems from the fact that BACBR-I also maximizes social welfare and restricts offers realizations according to their maximum volumes. Constraint (10) imposes that sell offer realization is integral. Next two constraints ensure that total bandwidth flow at particular link will not be greater than aggregated realizations of sell offers concerning this link, constraint (11) and that bandwidth flow at all links will be non-negative, constraint (12). Equation (12) is a flow conservation constraint that must be met for each end-to-end connection.

Comparing above models, both of them maximize social welfare on the basis of submitted offers that contain unit price and maximum volume. Both models treat bandwidth as indivisible good and support trading end-to-end connections. The BACBR-I and c-SeBiDA differs in the way of supporting end-to-end connection trading. Moreover, the BACBR-I is more comprehensive than c-SeBiDA, as it enables trading network resources consisting of modules with different size.

4. Experimental Studies

4.1. Test Instances

Experimental studies have been conducted on several test instances which are based on data from survivable network design library (SNDlib) available on the web site [11]. Although the SNDlib is a set of survivable fixed telecommunication network design problems, it provides information that is very important in bandwidth trading problems: network topology and a set of end-to-end connections. We consider three networks from SNDlib: sun, janos-us and giul39. Topologies of this networks are presented in Fig. 2. Table 1 contains the number of nodes, links and end-to-end connections for each considered network.

Table 1
Information about size of considered networks

Network	Nodes	Links	End-to-end connections
sun	27	102	67
janos-us	26	84	650
giul39	39	172	1471

Bandwidth allocation test instance besides aforementioned data requires specification of offers that are submitted for network resources. Offers have been generated according to the following rules:

– summarized bandwidth offered for sale (purchase) at link (end-to-end connection) equals the link capacity (end-to-end connection demand value) given by SNDlib;

– unit price of offer concerning link (end-to-end connection) is determined on the basis of the distance between nodes connected by this link (end-to-end connection) and some random factor that is used to differentiate prices of offers regarding the same link (end-to-end connection).

So also some other data from SNDlib such as nodes coordinates (used to calculate distance between nodes), links capacities, end-to-end connections demands come in useful for preparing test instances of bandwidth allocation problem.

For each network three offer variants have been generated with different average number of offers submitted for single link or end-to-end connection, respectively, 2, 4 and 6 offers per link or end-to-end connection. In all, nine test instances of bandwidth allocation problem have been prepared. All of them have been adjusted to the c-SeBiDA mechanism in which buy offers are submitted not for end-to-end connections, but for bundles of links. For each buy offer a sequence of links has to be specified that realizes suitable end-to-end connection. As a realization of end-to-end connection we choose randomly one of the three least expensive path realizations. Because end-to-end connection realization affects the social welfare obtained by the c-SeBiDA, we consider five variations of each test instance in which buy offers are submitted for different sequence of links realizing particular end-to-end connection. All test instances have been implemented in multicommodity market data model (M^3) [12].

4.2. Allocation Efficiency

The comparison of allocation efficiency obtained by both c-SeBiDA and BACBR-I mechanisms in all test instances is given in Table 2. Test instance is identified by network name and average number of offers submitted for single link or end-to-end connection. Table 2 does not contain the numerical values of social welfare achieved by both mechanisms, but only the relation between them assuming that social welfare determined by BACBR-I equals 1. Because in case of the c-SeBiDA mechanism we analyze five different variations of bundles of links generated for buy offers the maximium, average and minimum social welfare obtained by this mechanism in proportion to BACBR-I optimal allocation is presented.

Table 2
Comparison of c-SeBiDA and BACBR-I allocation efficiency

Network	Offers	BACBR-I	c-SeBiDA		
			max	avg.	min
sun	2	1	0.77	0.75	0.73
	4	1	0.77	0.72	0.65
	6	1	0.83	0.81	0.81
janos-us	2	1	0.84	0.8	0.77
	4	1	0.85	0.82	0.79
	6	1	0.86	0.84	0.81
giul39	2	1	0.8	0.79	0.79
	4	1	0.78	0.78	0.77
	6	1	0.82	0.81	0.81

Allocation efficiency of the c-SeBiDA is on average about 80% of BACBR-I allocation efficiency. In the best case social welfare achieved by c-SeBiDA accounts for 86% of social welfare determined BACBR-I. In case of network sun with 4 offers per single link or end-to-end connection on average the c-SeBiDA obtains only 65% of social welfare provided by BACBR-I.

4.3. Computational Efficiency

Table 3 presents information about number of variables and constraints of mathematical models related to particular mechanism and test instance identified by network name and average number of offers submitted for single link or end-to-end connection. The BACBR-I mathematical model has more variables and constraints than the c-SeBiDA model. The difference is substantial for the largest network giul39 with average 2 offers per link or end-to-end connection. In this test instance the BACBR-I model has about 80 and 10 times more variables and constraints, respectively, than the c-SeBiDA model.

Table 3
Number of variables (var.) and constraints (con.) in c-SeBiDA and BACBR-I mathematical models

Network	Offers	BACBR-I		c-SeBiDA	
		var.	con.	var.	con.
sun	2	7169	2581	335	772
	4	7515	3273	681	1464
	6	7864	3971	1030	2162
janos-us	2	56052	19888	1452	2988
	4	57606	22996	3006	6096
	6	59016	25816	4416	8916
giul39	2	256280	64077	3268	6708
	4	259589	70695	6577	13326
	6	262858	77233	9846	19864

The comparison of computational efficiency of c-SeBiDA and BACBR-I is given in Table 4. The table presents the time of solving mixed-integer linear programming prob-

Table 4
Comparison of c-SeBiDA and BACBR-I allocation time [s]

Network	Offers	BACBR-I	c-SeBiDA
sun	2	0.7	0.02
	4	0.89	0.02
	6	0.92	0.02
janos-us	2	15.8	0.05
	4	15.93	0.03
	6	14.19	0.07
giul39	2	499.01	0.05
	4	526.39	0.09
	6	512.71	0.17

lems related to considered test instances. Optimization has been performed by CPLEX 9.1 on computer with processor Intel Core2 Duo T8100 2.1 GHz, main memory 3 GB and 32-bit operating system MS Vista.

The BACBR-I mechanism requires more time than c-SeBiDA to determine optimal allocation. It is meaningful in case of the largest network giul39, for which BACBR-I model must find optimal allocation of links bandwidth for great number of end-to-end connections. Complexity of this task is reflected by the large number of variables and constraints of the BACBR-I model. It is worth noting that the allocation time of both mechanism is not rising a lot with increase of average number of offers submitted for single link or end-to-end connection.

5. Summary

This paper compares two single-round, socially efficient and double-sided auctions of indivisible network resources that represents different approaches of supporting end-to-end connection trading. The c-SeBiDA is a combinatorial auction that requires explicit single path specification posing realization of particular end-to-end connection. The BACBR-I enables submitting buy offers for pair of nodes that are the source and target of end-to-end connection. The former mechanism requires that buyer knows a network topology and chooses appropriate links. Explicit bundle of links specification in buy offers affects allocation efficiency of the c-SeBiDA which provides on average 20% less social welfare than the BACBR-I mechanism. The BACBR-I itself allocates the bandwidth links to the buyer assuring connections between selected nodes. It provides highest allocation efficiency, however, it requires more time to determine optimal allocation than the c-SeBiDA.

Acknowledgments

The authors acknowledge the Ministry of Science and Higher Education of Poland for partially supporting the research through project PBZ-MNiSW-02/II/2007.

References

[1] A. Iselt, A. Kirstadter, and R. Chahine, "Bandwidth trading – a business case for ASON?", in *Proc. 11th Int. Telecommun. Netw. Strat. Plann. Symp. Networks 2004*, Vienna, Austria, 2004, pp. 63–68.

[2] R. Rabbat and T. Hamada, "Revisiting bandwidth-on-demand enablers and challengers of a bandwidth market", in *Proc. Netw. Oper. Manage. Symp. NOMS 2006*, Vancouver, Canada, 2006, pp. 1–12.

[3] R. Jain and P. Varaiya, "Efficient market mechanisms for network resource allocation", in *Proc. Decis. Contr. Conf. CDC-ECC'05*, Seville, Spain, 2005, pp. 1056– 1061.

[4] P. Pałka, K. Kołtyś, E. Toczyłowski, and I. Żółtowska, "Model for balancing aggregated communication bandwidth resources", *J. Telecommun. Inform. Technol.*, no. 3, pp. 45–50, 2009.

[5] J. Abrache, T. G. Crainic, M. Gendreau, and M. Rekik, "Combinatorial auctions", *Ann. OR*, vol. 153, no. 1, pp. 131–164, 2007.

[6] M. Dramitinos, G. D. Stamoulis, and C. Courcoubetis, "An auction mechanism for allocating the bandwidth of networks to their users", *Comput. Netw.*, vol. 51, no. 18, pp. 4979–4996, 2007.

[7] R. Jain and J. Walrand, "An efficient mechanism for network bandwidth auction", in *Proc. Netw. Oper. Manage. Symp. NOMS 2008 Worksh.*, Salvador, Brazil, 2008, pp. 227–234.

[8] M. Pióro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. San Francisco: Morgan Kaufmann, 2004.

[9] G. Cheliotis, "Structure and dynamics of bandwidth markets", Ph.D. thesis, Athens, N.T.U., 2001.

[10] W. Stańczuk, J. Lubacz, and E. Toczyłowski, "Trading links and paths on a communication bandwidth markets", *J. Univers. Comput. Sci.*, vol. 14, no. 5, pp. 642–652, 2008.

[11] "Survivable network design library", 2009 [Online]. Available: http://sndlib.zib.de/

[12] "Open M3", 2008 [Online]. Available: http://www.openm3.org/

Kamil Kołtyś received the M.Sc. degree in computer science in 2007 from the Warsaw University of Technology, Poland. Currently he prepares his Ph.D. thesis in computer science at the Institute of Control and Computation Engineering at the Warsaw University of Technology. His research interests include decision support, optimization and bandwidth trading. His current research is focused on application of multicommodity turnover models to network resource allocation.

e-mail: K.J.Koltys@elka.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Izabela Żółtowska received the M.Sc. degree in 2000 and the Ph.D. degree in 2006 from the Warsaw University of Technology, Poland. She is an Assistant Professor at the Institute of Control and Computation Engineering at the Warsaw University of Technology. Her research focuses on the optimization models applied on the restructured competitive markets.

e-mail: I.Zoltowska@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Piotr Pałka, Eugeniusz Toczyłowski – for biography, see this issue, p. 60.

Paper

Price Method and Network Congestion Control

Krzysztof Malinowski, Ewa Niewiadomska-Szynkiewicz, and Przemysław Jaskóła

Abstract—Price instruments are useful in achieving market balance conditions in various markets. Those instruments can be also used for control of other composite systems. The formulation and basic properties of the Price Method are reviewed and then the congestion control by price instruments in a computer network is described and tested.

Keywords—computer network, congestion control, coordination strategy, price, price instruments, retail pricing problem.

1. Introduction

In this paper the objective is to present emerging, opportunities to use price mechanisms for management of computer networks. In recent years it was observed that the use of the price instruments could be made both to propose the new techniques for congestion control in data networks, in particular for Internet congestion control (ICC), and to better explain the existing congestion control mechanisms (the current TCP (transmission control protocol) congestion control protocols). It is useful to note at this point that Internet pricing is essential not only for better possible understanding of the network operation but it may also provide means to achieve rational behavior of the network users, who otherwise may "overgraze" the existing resources. Proper pricing of the network services is also necessary for many other purposes, not only for congestion control [1], [2]. Dynamic pricing may be useful, in particular, for balancing demand for access to application servers and for proper valuation of different classes and qualities of service [3].

2. Price Method; General Formulation and the Basic Facts

The Price Method, also known under the name of the Interaction Balance Method [4]–[7] represents a versatile approach to modeling, optimization and management of complex systems. This follows the natural role played by prices on various markets – to achieve the balance between the demand and the supply. The method allows for various problem formulations and price adjustment (price coordination) strategies. Problem formulation may either result from a formal partitioning (decomposition) of a large-scale optimization problem or it may result from more practical, application case-oriented, considerations. Price adjustment strategies can be derived from dual function optimization, from solving sets of the coordinating conditions or from practical possibilities existing in a particular application.

Consider the following, fairly general, deterministic problem related to complex system optimization:

$$\max_{x,u,y} \sum_i U_i(x_i, u_i, y_i)\,, \tag{1}$$

$$\text{subject to} \quad x_i \in I_i, \quad i = 1, \ldots, N\,, \tag{2}$$

$$y_i = F_i(x_i, u_i) \quad \text{and} \quad u_i = \sum_j H_{ij} y_j, \quad i, j = 1, \ldots, N\,, \tag{3}$$

$$\sum_i r_{li}(x_i) \le c_l, \quad l = 1, \ldots, L\,. \tag{4}$$

In the above formulation the objective is to maximize the aggregate utility across N involved entities (system elements); x_i is the vector of local decision variables (of dimension n_{xi}), u_i is the local (vector) interaction input (dim $u_i = n_{ui}$) and y_i (dim $y_i = n_{yi}$) is the local interaction output of the ith system element. In Eq. (3) it is assumed that the interaction output y_i is a unique deterministic function of x_i and u_i, and that the interaction coupling constraints $u_i = \sum_j H_{ij} y_j$ are linear. Obviously, these relations can be made more general. Similarly, local decision constraint sets I_i can be imposed also on u_i. Often these sets are in form of box constraints $I_i = [x_i^{\min}, x_i^{\max}]$ (with $x_i^{\min}$ being the vector of lower bounds and $x_i^{\max}$ the vector of upper bounds on the components of x_i). Constraint (4) is the global resource constraint; there are assumed to be L resources which may be required by the local entities; $r_{li}(x_i)$ represents the consumption of the lth resource by system element i. It is assumed that the optimization problem (1)–(4) has a solution.

The above optimization problem can be decomposed into N parallel local problems by introducing prices (in mathematical terms Lagrange multipliers) λ_i (dim $\lambda_i = n_{ui}$), $i = 1, \ldots, N$, associated with the coupling constraints $u_i = \sum_j H_{ij} y_j$, and μ_l, $l = 1, \ldots, L$, representing prices of the global resources. The ith local problem is then defined – for given prices – as follows:

$$\max_{x_i, u_i} \left[U_i(x_i, u_i, y_i) - \lambda_i^T u_i + \sum_{j=1}^{N} \lambda_j^T H_{ji} y_i - \sum_{l=1}^{L} \mu_l r_{li}(x_i) \right], \tag{5}$$

where $y_i = F_i(x_i, u_i)$ and subject to $x_i \in I_i$. Assuming unique solutions to the above N problems, $x_i(\lambda, \mu)$, $u_i(\lambda, \mu)$ and $y_i(\lambda, \mu)$, (where $y_i(\lambda, \mu) = F_i(x_i(\lambda, \mu), u_i(\lambda, \mu))$ and $\lambda^T = (\lambda_1^T, \ldots, \lambda_N^T)$, $\mu^T = (\mu_1^T, \ldots, \mu_L^T)$) one can seek such coordinating values λ^C and μ^C of λ and μ for which the coupling constraints (3) and the resource constraints (4) are satisfied, that is for $i = 1, \ldots, N$

$$y_i(\lambda^C, \mu^C) = F_i\left(x_i(\lambda^C, \mu^C), u_i(\lambda^C, \mu^C)\right),$$

$$u_i(\lambda^C, \mu^C) = \sum_j H_{ij} y_j(\lambda^C, \mu^C), \quad (6)$$

$$\sum_i r_{li}(x_i(\lambda^C, \mu^C)) \le c_l, \quad l = 1, \ldots, L, \quad (7)$$

$$\mu_l^C \ge 0, \quad \mu_l^C\left[\sum_i r_{li}(x_i(\lambda^C, p^C)) - c_l\right] = 0, \quad l = 1, \ldots, L. \quad (8)$$

The above conditions (7) and (8) result from the requirement to satisfy overall optimality conditions. It simple terms they state that the optimal prices of the resources must be nonnegative and that a positive price can be charged for the commonly available resource only when this resource is fully utilized, i.e., when the respective resource constraint is active. It should be observed that some or even all values of the components of λ_i^C may be negative. The assumption about the uniqueness property of the local problem solutions – for given price vectors μ and λ – is an essential one [4] for the conditions guaranteeing the existence of the coordinating prices λ^C, μ^C. If these solutions do not have this property, then the coordinating prices satisfying eqns. (6)–(8) may easily not exist. One could expect that the lack of uniqueness of $x_i(\lambda, \mu)$, $u_i(\lambda, \mu)$ for an isolated point (λ^0, μ^0), $\mu^0 \ge 0$ should not matter too much. Alas, it is easy to demonstrate with many important examples that if the local solutions are not unique for some pair (λ^0, μ^0), they are also not unique for any pair (λ, μ) that could satisfy the coordinating conditions. Now, the simple sufficient conditions guaranteeing the uniqueness of $x_i(\lambda, \mu)$, $u_i(\lambda, \mu)$ (and hence of $y_i(\lambda, \mu)$) for any feasible pair (λ, μ) are: the strict concavity of the function $U_i(x_i, u_i, y_i)$ with respect to its arguments, linearity of F_i and the convexity of r_{li} and of the set I_i. More general conditions can also be given [5], [8] but in a general nonlinear case the desirable uniqueness property of the local problem solutions is, unfortunately, not easily achievable. Assuming the existence of λ^C, μ^C, one can propose a number of algorithms for iterating the values of λ, μ. The basic strategy is to use the following gradient method:

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} + \gamma\left[u_i(\lambda^{(k)}, \mu^{(k)}) - \sum_j H_{ij} y_j(\lambda^{(k)}, \mu^{(k)})\right], \quad (9)$$

$$\mu_l^{(k+1)} = \left\{\mu_l^{(k)} + \gamma\left[\sum_i r_{li}(x_i(\lambda^{(k)}, \mu^{(k)})) - c_l\right]\right\}_+ \quad (10)$$

for $i = 1, \ldots, N$, $l = 1, \ldots, L$, where γ is a positive step size, and k is the iteration index. The algorithm defined by (9) and (10) is a gradient strategy – with price projection on the feasible range – since, assuming the uniqueness of $x_i(\lambda, \mu)$, $u_i(\lambda, \mu)$ for every feasible pair (λ, μ) and each i, the expressions in square brackets in (9), (10) are the slopes (with minus sign) of the dual function defined as the sum of the maximum local performance values Eq. (5) minus the term $\sum_l \mu_l^{(k)} c_l$. The dual function attains its minimum at λ^C, μ^C.

The above price adjustment strategy will converge for sufficiently small value of γ if all local utility functions are strongly concave, functions F_i are linear, functions r_{li} are convex, sets I_i are convex and there exists a feasible point satisfying all the constraints, such that all the inequality constraints are inactive. Algorithm (9) and (10) has a distributed character. In particular, the adjustments of the resource prices μ_l can be performed for each price independently from the other price adjustments. This property appears to be most useful.

3. Network congestion control and price-based schemes

With Internet, the Price Method has found a large-scale system for which this method seems to be quite well suited. In several congestion control mechanisms, as recently proposed [9]–[14], the network is represented by S traffic sources, representing particular source-destination pairs, and a grid of a set of L links. The links, together with associated routers, are the network resources, of limited traffic carrying capacity c_l. Each source i is supposed to use a set $L(i) \subseteq L$ of links, and has at time k an associated transmission rate x_i; the set of transmission rates determines the aggregate flow $y_l(k)$ through each link, by the equation:

$$y_l(k) = \sum_{i \in S(l)} x_i(k), \quad (11)$$

where $S(l)$ is the set of all sources transmitting through the link l. Then, the feedback mechanism communicates to sources the congestion information about the network. This congestion measure – the price $p_l(k)$ – is a positive valued quantity associated with link l. The fundamental assumption is made that sources have access to the aggregate price ($q_i(k)$, $i = 1, \ldots, S$) of all links in their route:

$$q_i(k) = \sum_{l \in L(i)} p_l(k). \quad (12)$$

4. Network Control by Price Instruments, Application of Price Method

Let us consider the network with S sources and L links. Assume now that the traffic sources (or source-destination pairs) are indeed utility oriented, i.e., they have utilities $U_i(x_i)$ expressed in monetary terms and are willing to maximize profits equal to utility minus payment charged by the network. Then, the equilibrium rate x_i^{se} will solve the following local source problem:

$$\max_{x_i}[U_i(x_i) - q_i^{se} x_i], \quad (13)$$

where $q_i^{se} = \sum_{l \in L(i)} p_l^{se}$ and p^{se} denotes vector of equilibrium prices. The role of prices p is to coordinate the actions of the individual sources; in fact to ensure that solutions of (13) together solve the network flow optimization problem

$$\max_{x\geq 0}\sum_i U_i(x_i) \quad \text{subject to} \quad \sum_{i\in S(l)} x_i \leq c_l, \quad l=1,\ldots,L, \tag{14}$$

where c_l is the lth link capacity. This is particular, simple instant of problem (1)–(4). Hence, it would seem possible – at least in theory – to propose the following congestion control scheme: for given link prices $p_l(k)$, $l=1,\ldots,L$, at time k, the sources solve local problems

$$\max_{x_i\in I_i}[U_i(x_i) - q_i(k)x_i], \tag{15}$$

where $I_i = [x_i^{\min}, x_i^{\max}]$. The solutions $x_i^s(k) = x_i(q_i(k))$ are signaled to all concerned links, which then adjust their link prices – for the next iteration $k+1$ – according to the following rule

$$p_l^{(k+1)} = \left\{ p_l^{(k)} + \gamma\left[\sum_{i\in S(l)} x_i^s(k) - c_l\right]\right\}_+ \tag{16}$$

and is a positive step – chosen to allow the scheme to converge. Then the new link prices are signaled to links, etc., until the convergence is obtained.

5. Dynamic Routing

Although it would require a serious modification of Internet protocols to implement the described scheme of coordination by price instruments in a distributed manner, this algorithm can be instead used by the centralized controller, so called bandwidth broker. One can observe, that further improvement of the profit can be achieved by conscious choice of routing, in the systems with redundant paths existing between nodes. Here is an additional profit of using price instruments for network coordination. Equilibrium prices p_l, calculated in the process of network balancing can be interpreted as link congestion indicators. As such they can be used by external shortest-path routing algorithm.

There is a problem connected with this approach. The calculation of routes based on dynamic metrices leads to the change in the load pattern, and necessity of re-calculating allocations, and so on. In other words it may lead to oscillations, what was observed in experiments. To avoid this problem, time scales of routing and bandwidth allocation have to be separated. For example, rerouting can be performed every few minutes.

The following rerouting scheme is proposed:

Step 0: Sort the list of active sources in descending order w.r.t. $x_i^{\max}$.

Step 1: Choose the first element from the sorted list and find a shortest length route for it using Dijkstra algorithm, basing on dynamic metrices defined in Eq. (17); remove this source from the list.

Step 2: Calculate the new equilibrium prices pattern of the network solving the problem (15)–(16), update the metrices.

Step 3: Proceed until the route is found for every pair source-destination.

The critical issue is to propose an additive metric for the shortest length search. It must be a scalar value, bound to the link, which incorporates information about the path load and capacity at the same time. Finally, the proposed dynamic metric d_l includes the congestion prices, and takes the form:

$$d_l = p_l + \alpha\beta e^{-\frac{c_l}{c^{\max}}}, \qquad \alpha \in <0,1>, \tag{17}$$

where $c^{\max}$ is the highest link capacity in the network, $\beta = \max_{i\in S} x_i^{\max}$. Equation (17) takes into account, that at the initial iterations the network is "empty", and none links are congested – thus $p_l = 0$. The exponential term reflects the fact, that capacity is a non-additive metric, that must be dealt with by shortest-path algorithm. In this case the routes with the highest capacities are occupied in the first place by the source-destination pairs with the highest transmission rates.

6. Results of Experiments

The simulation experiments were performed for a simple network configuration. Six sources were connected to the network presented in Fig. 1. All were assumed to trans-

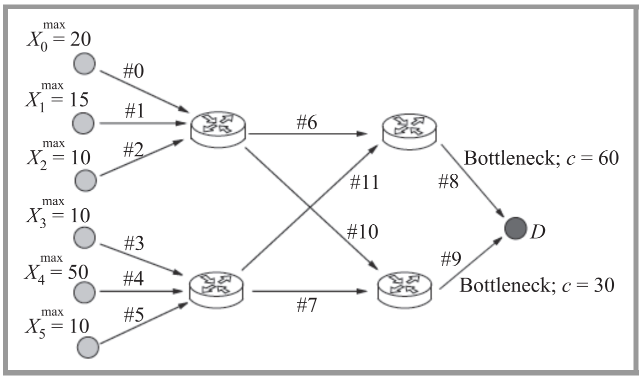


Fig. 1. Test network.

mit to the same destination, symbolized with a dark circle. The sources compete for the bandwidth of two bottlenecks (link 8 and link 9). Two algorithms for the bandwidth broker were compared, one that gives an optimal solution of the bandwidth allocation problem – mixed integer programming (MIP), and price-based scheme described in the previous section. The aggregate utilities were calculated.

The response of bandwidth broker is presented in Table 1. Both the MIP and price-based algorithm gave the same result in – value of the utility function in both cases was 1659.4. This result was obtained under the assumption that all sources in the initial stage were sorted. For comparison in Table 1 (right-hand side) the performance of a modified price-based algorithm is presented, in which

Table 1
Links and rates, MIP, and price-based algorithm with sorting (left), price-based algorithm without sorting (right)

	With sorting		Without sorting	
s	link numbers	rate (x)	link' numbers	rate' (x)
0	6; 8	15.00	6; 8	19.00
1	9; 10	11.25	6; 8	14.00
2	9; 10	6.25	6; 8	9.00
3	7; 9	6.25	8; 11	9.00
4	8; 11	45.00	9; 7	30.00
5	7; 9	6.25	8; 11	9.00
$U(x)$	1659.4		1510.0	

only the Lagrange multipliers are used to represent link costs (metrics). No sorting was performed in the initial stage in this case. The obtained value of the utility function was 1510.0.

7. Final Remarks

In this paper the objective was to present various selected approaches to pricing, in particular concerned with the classical role of prices in balancing markets and systems in view of limited resources. The formulation and the main properties of the Price Method were reviewed. This method, developed with various modifications in the 1970-ties, still receives much attention. To illustrate this point possible application of the price-based mechanisms to Internet congestion control was discussed. In fact the size and the problems with operation of the Internet motivate the renewed interest in those mechanisms. The pricing "technology" is very much needed to control modern communications because there do not exist other optimization based technologies that could cope with the challenges offered by the Network.

Acknowledgments

The paper was partially supported by Polish Ministry of Science and Higher Education grant N N514 416934.

References

[1] L. A. DaSilva, "Pricing for QoS-enabled networks: a survey", *IEEE Commun. Surv.*, 2000 [Online]. Available: http://www.comsoc.org/pubs/surveys

[2] M. Falkner, M. Devetsikiotis, and I. Lambadaris, "An overview of pricing concepts for broadband IP networks", IEEE Commun. Surv., 2000 [Online]. Available: http://www. comsoc.org/pubs/surveys

[3] G. Armitage, *Quality of Service in IP Networks: Foundations for a Multi-Service Internet*. Indianapolis: Macmillan, 2000.

[4] W. Findeisen, F. N. Bailey, M. Brdys, K. Malinowski, P. Tatjewski, and A. Wozniak, *Control and Coordination in Hierarchical Systems*. London: Wiley, 1980.

[5] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: Macmillan, 1970.

[6] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical, Multilevel, Systems*. New York: Academic Press, 1970.

[7] *Large-Scale Systems Control and Decision Making*, H. Tamura and T. Yoshikawa, Eds. New York: Marcel Dekker, 1990.

[8] K. Malinowski, "Applicability of the Lagrange multiplier method to optimization of quadratic systems", *IEEE Trans. Autom. Contr.*, vol. AC-22, no. 4, pp. 648–651, 1977.

[9] X-R. Cao and H-X. Shen, "Internet pricing: comparison and examples", in *39th IEEE Conf. Decis. Contr.*, Sydney, Australia, 2000.

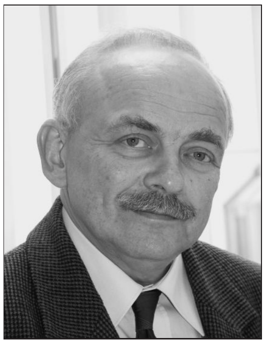
[10] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability", *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[11] R. J. La and V. Anantharam, "Charge-sensitive TCP and rate control in the Internet", Rep., 2000 [Online]. Available: http://www.eecs.berkeley.edu

[12] S. H. Low and D. E. Lapsley, "Optimization flow control. I: Basic algorithm and convergence", *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, 1999.

[13] K. Malinowski, "Optimization network flow control and price coordination with feedback; proposal of a new distributed algorithm", *Comput. Commun.*, vol. 25, pp. 1028–1036, 2002.

[14] I. C. Paschalidis and J. N. Tsitsiklis, "Congestion-dependent pricing of network services", *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 171–184, 2000.



Krzysztof Malinowski received his Ph.D. in 1974, D.Sc. in 1978 from the Warsaw University of Technology (WUT), Poland. He works as a Professor of control and information engineering at the same University. He is the Head of the Control and Systems Division. He is also a Professor at the Research and Academic Computer Network (NASK) and a member of the Polish Academy of Sciences. He is the author or co-author of four books and over 130 journal and conference papers. For many years he was involved in research on hierarchical control and management methods. His current interests are control, simulation and optimization of communication networks. He was a visiting Professor at the University of Minnesota in 1976 and in 1978/1979, and then, from 1981 till 2001, served as a consultant to the Decision Technologies Group of UMIST in Manchester, UK. He was the Director of Institute of Control and Computation Engineering in 1984–1996 and Dean of the Faculty of Electronics and Information Technology in 1996–1999. From 2003 to 2009 he was the Director for Research of NASK. From 2000 to 2002 he was the project manager, on the Polish side, of the QOSIPS project in the IST Programme of the FP 5. The QOSIPS project was concerned with development of the system consisting of the Quality Module and the Pricing Module, for tariff establishment and monitoring of telecommunication services of differentiated quality of service.

e-mail: K.Malinowski@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

e-mail: Krzysztof.Malinowski@nask.pl
Research Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland

Ewa Niewiadomska-Szynkiewicz received her Ph.D. in 1996, D.Sc. in 2006 from the Warsaw University of Technology. She works as a Professor of control and computation engineering at the Warsaw University of Technology. She is the Head of the Complex Systems Group. She is also an associate professor at the Research and Academic Computer Network (NASK), and the Director for Research of NASK since 2009. She is the author or co-author of three books and over 90 journal and conference papers. Her research interests focus on complex systems modeling and control, computer simulation, global optimization, parallel calculations and computer networks. She was involved in a number of research projects including three EU projects, coordinated the Groups activities, managed organization of a number of national-level and international conferences.

e-mail: ens@ia.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

e-mail: ewan@nask.pl
Research Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland

Przemysław Jaskóła received his M.Sc. in computer science from the Warsaw University of Technology, Poland, in 1999. Currently he is a Ph.D. student in the Institute of Control and Computation Engineering at the Warsaw University of Technology. Since 2005 he works for Research and Academic Computer Network (NASK). His research area focuses on hierarchical optimization and computer networks.

e-mail: Przemyslaw.Jaskola@nask.pl
Research Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland

Paper

Computational Methods for Two-Level 0-1 Programming Problems through Distributed Genetic Algorithms

Keiichi Niwa, Tomohiro Hayashida, and Masatoshi Sakawa

Abstract—In this paper, we consider a two-level 0-1 programming problem in which there is not coordination between the decision maker (DM) at the upper level and the decision maker at the lower level. We propose a revised computational method that solves problems related to computational methods for obtaining the Stackelberg solution. Specifically, in order to improve the computational accuracy of approximate Stakelberg solutions and shorten the computational time of a computational method implementing a genetic algorithm (GA) proposed by the authors, a distributed genetic algorithm is introduced with respect to the upper level GA, which handles decision variables for the upper level DM. Parallelization of the lower level GA is also performed along with parallelization of the upper level GA. The proposed algorithm is also improved in order to eliminate unnecessary computation during operation of the lower level GA, which handles decision variables for the lower level DM. In order to verify the effectiveness of the proposed method, we propose comparisons with existing methods by performing numerical experiments to verify both the accuracy of the solution and the time required for the computation.

Keywords—distributed genetic algorithm, Stackelberg solution, two-level 0-1 programming problem.

1. Introduction

In the real world, we can often encounter situations that there are multiple decision makers (DMs) in hierarchically structured organizations, and decisions may be taken serially or simultaneously in order to optimize each of the objectives. This kind of problem has been formulated as a two-level programming problem [1]. In two-level programming problems, the upper level DM makes his/her decision first, and then, with full knowledge of the decision of the upper level DM, the lower level DM makes his/her decision in order to optimize his/her own objective function. According to this rule, the upper level DM also makes a decision so as to optimize the objective function of self. The solution defined as the above mentioned procedure is a Stackelberg solution. In this paper, both the upper level and the lower level have one DM, and the problem is treated as a two-level 0-1 programming problem in which both DMs treat all of their decision variables as 0-1 variables.

As an overview of research dealing with two-level programming problems that include discrete variables, Bard *et al.* presented an algorithm based on the branch-and-bound approach in order to derive the Stackelberg solution for two-level 0-1 programming problems [2] and two-level mixed integer programming problems [3]. Wen *et al.* [4] have presented a computation method for obtaining the Stackelberg solution to two-level programming problems which have 0-1 variables for the decision variables in the upper level and real variables for the decision variables in the lower level.

On the other hand, the adaptive process of systems in the natural world has been explained, and genetic algorithms (GAs) which imitate the evolution occurring in living organisms have been receiving attention at international conferences related to GAs, publications by Goldberg [5], as have methodologies for optimization, adaptation and learning. GAs have also been adopted for a variety of combinatorial optimization problems, and their effectiveness has been reported [6].

An example of research related to two-level programming problems using GAs is given by Anandalingam, *et al.* [7] which presents a method for deriving a Stackelberg solution for two-level linear programming problems. Also, Nishizaki, *et al.* presented an algorithm based on GAs in order to derive the Stackelberg solution for two-level integer programming problems [8] and two-level mixed integer programming problems [9]. In order to derive a Stackelberg solution for 0-1 programming problems related to two-level decentralized systems, the authors [10] have also proposed a computational method that adopts the double string proposed by Sakawa, et al as the individual representation. In order to improve the computational accuracy of approximate Stakelberg solutions, the authors have proposed computational methods that implement sharing [11] and cluster analysis [12] methods. Furthermore, the authors have proposed a computational method using parallel genetic algorithm [13]. Use of these methods allows for the derivation of approximate Stackelberg solutions with relatively high precision and in a relatively short time, but there is still room for improvement, particularly with regards to calculation times.

Therefore, this paper focuses on two-level 0-1 programming problems, and proposes an improved computational method that addresses problems related to the computational method proposed by the authors for deriving the Stackelberg solution. Specifically, a distributed genetic algorithm is introduced with respect to the upper level GA, which handles decision variables for the upper level DM, in order to improve the computational accuracy of approximate Stackelberg solutions and decrease the computational time of a computational method implementing a genetic algorithm proposed by the authors. Also, parallelization of the lower level GA is performed along with parallelization of the upper level GA. The proposed algorithm is also improved in order to eliminate unnecessary computation during operation of the lower level GA, which handles decision variables for the lower level DM. In order to verify the effectiveness of the proposed method, we propose comparisons with the existing method and the computational method using parallel GA by performing numerical experiments to verify both the accuracy of the solution and the time required for the computation.

2. Two-Level 0-1 Programming Problem

For the sake of brevity, we denote the upper and lower level DMs by DM1 and DM2, respectively. The two-level 0-1 programming problem is expressed as

$$\left.\begin{array}{ll} \underset{\mathbf{x}}{\text{maximize}} & z_1(\mathbf{x},\mathbf{y}) = \mathbf{c}_1\mathbf{x} + \mathbf{d}_1\mathbf{y} \\ \text{where } \mathbf{y} \text{ solves} & \\ \underset{\mathbf{y}}{\text{maximize}} & z_2(\mathbf{x},\mathbf{y}) = \mathbf{c}_2\mathbf{x} + \mathbf{d}_2\mathbf{y} \\ \text{subject to} & A\mathbf{x} + B\mathbf{y} \leqq \mathbf{b} \\ & \mathbf{x} \in \{0,1\}^{n_1}, \mathbf{y} \in \{0,1\}^{n_2}, \end{array}\right\} \tag{1}$$

where $\mathbf{x} = (x_1,\ldots,x_{n_1})^T$, and $\mathbf{y} = (y_1,\ldots,y_{n_2})^T$ are the vectors of decision variables for DM1 and DM2; $z_1(\mathbf{x},\mathbf{y})$, and $z_2(\mathbf{x},\mathbf{y})$ respectively represent the objective functions of DM1 and DM2; $\mathbf{c}_1 = (c_{11},\ldots,c_{1n_1})$, $\mathbf{d}_1 = (d_{11},\ldots,d_{1n_2})$, $\mathbf{c}_2 = (c_{21},\ldots,c_{2n_1})$, and $\mathbf{d}_2 = (d_{21},\ldots,d_{2n_2})$ denote the coefficient vectors of the objective functions; A and B are $m \times n_1$ and $m \times n_2$ coefficient matrices in the constraints, respectively; $\mathbf{b} = (b_1,\ldots,b_m)^T$ is a coefficient vector of the right hand side of the constraints; the superscript T means transposition of a vector.

For the sake of simplicity, in this paper, it is assumed that each component of A, B, $\mathbf{b}$, $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{d}_1$, and $\mathbf{d}_2$ is positive.

It is possible to express the process for choosing the Stackelberg solution for a two-level 0-1 programming problem in the following manner. Each decision maker completely knows objective functions and constraints of the opponent and self, and DM1 first makes a decision and then DM2 makes a decision in order to maximize the objective function with full knowledge of the decision of DM1. That is to say, when the decision by DM1 is denoted $\hat{\mathbf{x}}$, DM2 solves the 0-1 programming problem (2) with parameters $\hat{\mathbf{x}}$, choosing the optimal solution $\mathbf{y}(\hat{\mathbf{x}})$ as the rational reaction to $\hat{\mathbf{x}}$.

$$\left.\begin{array}{ll} \underset{\mathbf{y}}{\text{maximize}} & z_2(\hat{\mathbf{x}},\mathbf{y}) = \mathbf{d}_2\mathbf{y} + \mathbf{c}_2\hat{\mathbf{x}} \\ \text{subject to} & B\mathbf{y} \leqq \mathbf{b} - A\hat{\mathbf{x}} \\ & \mathbf{y} \in \{0,1\}^{n_2} \end{array}\right\} \tag{2}$$

Under this premise, DM1 also determines $\mathbf{x}$ by choosing the value which maximizes its own objective function. For problems which adopt the Stackelberg solution to conceptualize their solution, it is assumed that there is no consensus among DMs that might mutually constrain decisions. Putting it another way, their relationship may be described as non-cooperative.

3. GA Based Computational Method

We propose a computational method through GA in order to obtain Stackelberg solutions to the two-level 0-1 programming problems. In Subsections 3.1, 3.2 and 3.3, first, we describe fundamental elements of GA, which are coding procedure, a decoding procedure and genetic operators, used in the computational method using GA [10]. In this paper, we call the computational method using GA [10] normal GA (NGA). Furthermore, we show additional elements used in the proposed computational method using distributed genetic algorithm and the computational method using parallel genetic algorithm [13] in Subsections 3.4 and 3.5. Finally, the algorithm used in the proposed computational method using distributed genetic algorithm is described in Subsection 3.6.

3.1. Coding and decoding

When solving 0-1 programming problems using GAs, binary strings are usually adopted to express individuals [5], [14]. However, under this representation it is possible that infeasible individuals that do not satisfy the constraints may be generated, so there is a danger that the performance of the GAs may degrade. Thus, in this paper, a double string [6] is used which is composed of the substring corresponding to the decision of DM1, $\mathbf{x}$, and the substring corresponding to the decision of DM2, $\mathbf{y}$, as shown in Fig.1 in order to derive only feasible solutions. The decisions of DM1 and DM2 are handled by performing genetic operators on each sub-individual. We call the GA operating to the decision $\mathbf{x}$ of DM1 the upper level GA, and the GA operating to the decision $\mathbf{y}$ of DM2 the lower level GA.

← Individual for $\mathbf{x}$ →			← Individual for $\mathbf{y}$ →		
$i_x(1)$	$\cdots$	$i_x(n_1)$	$i_y(1)$	$\cdots$	$i_y(n_2)$
$S_{i_x(1)}$	$\cdots$	$S_{i_x(n_1)}$	$S_{i_y(1)}$	$\cdots$	$S_{i_y(n_2)}$

Fig. 1. Double string.

In Fig. 1 $s_{i_x(m)} \in \{0,1\}, i_x(m) \in \{1,\ldots,n_1\}$, and for $m \neq m'$ it is assumed that $i_x(m) \neq i_x(m')$. Similarly, $s_{i_y(m)} \in \{0,1\}, i_y(m) \in \{1,\ldots,n_2\}$, and for $m \neq m'$ it is assumed that $i_y(m) \neq i_y(m')$. Also, in the double string, $i_x(m)$, $i_y(m)$ and $s_{i_x(m)}$, $s_{i_y(m)}$ express indexes of the elements of each solution vector respectively, and their values.

In order to generate only feasible solutions, a decoding algorithm proposed by the authors [10] is also applied to the upper level and the lower level GA.

3.2. Reproduction

We describe the reproduction operator of the lower level GA. Substituting the given value $\mathbf{x}$ of the decision variable in the upper level GA and the value of $\mathbf{y}$ obtained by decoding individuals in the lower level GA into the objective function of DM2, $z_2(\mathbf{x},\mathbf{y})$, the value of the evaluation function for each individual is obtained. Next, the fitness value for each individual is derived using linear scaling, and the individuals remaining in the next generation are determined by applying elitist expected value selection.

We describe the reproduction operator of the upper level GA. Substituting the value of $\mathbf{x}$ obtained by decoding individual in the upper level GA and the value of the rational reaction $\mathbf{y}(\mathbf{x})$ obtained by applying the lower level GA into the objective function of DM1, $z_1(\mathbf{x},\mathbf{y}(\mathbf{x}))$, the value of the evaluation function for each individual is obtained. Next, the fitness value for each individual is calculated by applying linear scaling and adopting a clustering method. The individuals remaining in the next generation are determined by applying elitist expected value selection based on these fitness values.

3.3. Crossover and Mutation

For double strings, if single-point or multi-point crossover operators are performed then there is a possibility that infeasible individuals may be generated because the indexes occurring in the offspring, $i_x(m)$, $i_x(m')$, $m \neq m'$ or $i_y(m)$, $i_y(m')$, $m \neq m'$, may have the same number. When solving the traveling salesman problem or the scheduling problem through GAs, this kind of violation occurs. In order to circumvent such violation, partially matched crossovers (PMX) have been devised. In this paper, a modified version of PMX is used in order to handle the double strings proposed by Sakawa *et al.* [6]. Also, when determining whether or not to apply the crossover operator, a probability p_c is used. Its value is set in advance.

PMX procedure

Step 1: For two individuals expressed using double strings, s_1 and s_2, two crossover points are set at random.

Step 2: According to PMX, the upper strings of s_1 and s_2, along with the corresponding lower strings are reordered, generating s_1' and s_2'.

Step 3: For double strings, the offsprings, s_1'' and s_2'', resulting from the application of the revised PMX are obtained by exchanging the lower strings between the two crossover points s_1' and s_2'.

It is well recognized that the mutation operator plays a role of local random search in genetic algorithms. In this paper, the mutation operator is applied to each string, and inversion is used for index strings. For binary strings, mutation of bit-reverse type is adopted. When applying the mutation operator to individuals, it is first determined whether or not the mutation operator will be applied to an individual according to the mutation probability p_m. In the case that mutation is applied, it is then determined whether to apply inversion or bit-reverse according to the mutation selection constant M_{Pum}.

Mutation procedure

Step 1: For an individual s, expressed using a double string, a random number r_m is generated. If $r_m \leqq M_{Pum}$, a point on the 0-1 string is chosen at random and bit-reverse is performed, yielding s_1'. Otherwise, Step 2 is adopted.

Step 2: Two points on the index string are chosen at random, and inversion is applied to the substring between the two points, yielding s_2'.

3.4. Application of the parallel genetic algorithm

In genetic algorithms, it is possible to perform parallel processing in the greater part of the operations included in the algorithm. In reproduction operations, however, because it is necessary to calculate evaluation values for each individual in a population, and based on that value determine the fitness of each individual, direct application of parallel processing is difficult. Research related to the parallelization of GAs started with improvements to such barriers to the implementation of parallelization, and a variety of types of models have been proposed and their effectiveness noted by numerous researchers [15], [16], [17]. Today, GAs that implement parallel processing have come to be called parallel genetic algorithms.

The computational method proposed by the authors [13] divides the lower level GA operations and assigns them across multiple processors. Also, the computational method adopts the single-population master-slave GAs as the upper level GA. By assigning the calculation of individual fitness values, crossover operator and mutation operator to multiple processors, calculation times are reduced. We call the computational method proposed by the authors the computational method using parallel GA (PGA). However, while the computational method using parallel GA succeeds in obtaining good approximate solutions and reducing the amount of computational time, there is still likely much more room for improvement. In this study, therefore, we aim for further improvements of the precision of approximate solutions and further reductions in computational time, and consider parallelization of the upper level GA and the lower level GA implemented by the computational method using parallel GA.

The multiple-population genetic algorithms performs parallel processing by dividing the population and assigning the partial populations (sub-populations) to multiple processors. If we adopt this model as our computational method, it is possible to use the computational method using GA proposed by authors [12] without a lot of modifications. In that case, the multiple-population genetic algorithm is adopted as the upper level GA. The upper level GA operations proposed by authors are applied for partial population assigned to each processor. Additionally, migration operator is performed in every migration interval. Also, it is able to divide the lower level GA operations and assign them across multiple processor in the same way as PGA. In this paper, we employ a multiple-population genetic algorithms (distributed genetic algorithms) so as to obtain good approximate Stackelberg solutions and reduce calculation times. We call the proposed computational method using distributed GA (DGA).

3.5. Lower Level GA Avoidance Procedures

In this paper we introduce a storage region as shown in Fig. 2.

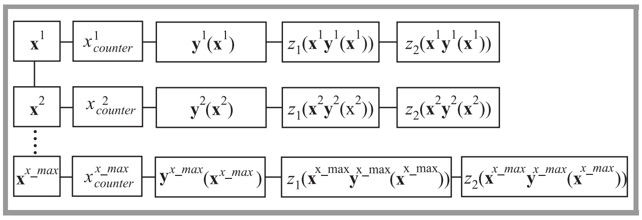


Fig. 2. Storage for saving $\mathbf{x}$ and $\mathbf{y}(\mathbf{x})$.

Here, $\mathbf{x}^i, i = 1,\ldots,x\_max$ indicate those values of $\mathbf{x}$ that were used in the past for handling individuals of the upper level GA, and $\mathbf{y}^i(\mathbf{x}^i), i = 1,\ldots,x\_max$ indicate the values of the rational reactions associated with $\mathbf{x}^i$ obtained by the lower level GA. $x^i_{counter} \in \{1,2,\ldots,y\_max\}, i = 1,\ldots,x\_max$ indicate the number of times that the lower level GA was used to find the rational reaction $\mathbf{y}^i(\mathbf{x}^i)$ for $\mathbf{x}^i$. x_max indicates the maximum number of DM1 decisions $\mathbf{x}$ saved, and y_max indicates the maximum number of times that the lower level GA can be repeatedly used to find the rational reaction $\mathbf{y}^i(\mathbf{x}^i)$ for $\mathbf{x}^i$. $z_1(\mathbf{x}^i,\mathbf{y}^i(\mathbf{x}^i))$ and $z_2(\mathbf{x}^i,\mathbf{y}^i(\mathbf{x}^i))$ are stored $\mathbf{x}^i,\mathbf{y}^i(\mathbf{x}^i)$ values used in place of DM1 and DM2 objective functions. By using the following algorithm, the number of applications of the lower level GA is reduced, and unnecessary calculation times eliminated.

Storage of the rational reaction $\mathbf{y}(\mathbf{x})$ and lower level GA avoidance procedures

Step 1: If there exists in $\mathbf{x}^i$ an upper level GA individual $\bar{\mathbf{x}}$, proceed to Step 2. If one does not exist, then check if the number of $\mathbf{x}^i$ has reached x_max, and if so continue on to Step 3. If not, proceed to Step 4.

Step 2: If $x^i_{counter}$ has reached y_max, then the saved $\mathbf{y}^i(\mathbf{x}^i)$ is returned to the upper level GA as the rational reaction and the algorithm terminates. If not reached, proceed to Step 4.

Step 3: Select the least of the values $z_1(\mathbf{x}^i,\mathbf{y}^i(\mathbf{x}^i))$ from the saved $\mathbf{x}^i$, and take that $\mathbf{x}^i$ value as $\mathbf{x}^k$. After applying the lower level GA and thus obtaining the rational reaction $\mathbf{y}(\bar{\mathbf{x}})$ for $\bar{\mathbf{x}}$, if $z_1(\mathbf{x}^k,\mathbf{y}^k(\mathbf{x}^k)) \leq z_1(\bar{\mathbf{x}},\mathbf{y}(\bar{\mathbf{x}}))$, save $\bar{\mathbf{x}}$, $\mathbf{y}(\bar{\mathbf{x}})$, $z_1(\bar{\mathbf{x}},\mathbf{y}(\bar{\mathbf{x}}))$, $z_2(\bar{\mathbf{x}},\mathbf{y}(\bar{\mathbf{x}}))$ in the storage region $\mathbf{x}^k$, and terminate the algorithm.

Step 4: After obtaining the rational reaction $\mathbf{y}(\bar{\mathbf{x}})$ for $\bar{\mathbf{x}}$ by applying the lower level GA, save $\bar{\mathbf{x}}$, $\mathbf{y}(\bar{\mathbf{x}})$, $z_1(\bar{\mathbf{x}},\mathbf{y}(\bar{\mathbf{x}}))$, $z_2(\bar{\mathbf{x}},\mathbf{y}(\bar{\mathbf{x}}))$, and terminate the algorithm.

Implementation of the algorithm described above improved upon previous methods.

3.6. The Algorithm for the Improved Computational Method

The algorithm used in the computational method after improvement can be described as follows, N_p denotes the number of processors.

Step 1: For each processor $q, q = 1,\ldots,N_p$, apply the upper level GA operations on Step1 through Step 7. Taking the generation of the upper level GA as $t_{uq} := 0$, N_u initial individuals are randomly generated.

Step 2: For each individual $\mathbf{x}$ in the upper level GA, determine whether or not to apply the lower level GA, and find the number of lower level GA to apply, N_{ul}. For those individuals N_{ul}, apply the lower level GA operations in Step 2-1 through Step 2-3, and obtain the rational reaction $\mathbf{y}(\mathbf{x})$. For those $N_u - N_{ul}$ individuals to which the lower level GA will not be applied, take the saved $\mathbf{y}(\mathbf{x})$ as the rational reaction, and proceed to Step 4.

Step 2-1: Set $t_l := 0$. Randomly generate N_l lower level GA individuals y, and take these as the initial population of the lower level GA . Proceed to Step 2-2.

Step 2-2: By using $\mathbf{x}$ given as the upper level GA individual and $\mathbf{y}$ generated by the lower level GA, the DM2 objective function value is calculated. After applying linear scaling to the value, the reproduction operator is applied. Proceed to Step 2-3.

Step 2-3: If t_l has exceeded the previously defined a maximum number of generation M_l, take the individual with the best fitness value as the optimal individual $\mathbf{y}(\mathbf{x})$, and proceed to Step 3. Otherwise, apply crossover operator and mutation operator to each lower level GA individual, let $t_l = t_l + 1$, and proceed to Step 2-2.

Step 3: By using the lower level rational reaction $\mathbf{y}(\mathbf{x})$ obtained by operation of the lower level GA and the individual $\mathbf{x}$ of the upper level GA, calculate the values for the DM1 and DM2 objective functions. Perform the procedures required to save $\mathbf{x}$ and its rational reactions $\mathbf{y}(\mathbf{x})$ to the storage region, and proceed to Step 4.

Step 4: Calculate the DM1 objective function for each upper level GA individual $\mathbf{x}$, and after performing linear scaling, apply the clustering method to measure the level of convergence of the individuals. Depending upon the degree of convergence, calculate the fitness value of each individual. Proceed to Step 5.

Step 5: If t_{uq} has exceeded the previously set a maximum number of generation M_u, then terminate the algorithm. In that case, the individual obtained up to that generation with the best fitness value is taken as the optimal individual $(\mathbf{x},\mathbf{y})$. Otherwise, proceed to Step 6.

Step 6: Reproduction operator is performed using the fitness values of each individual of the upper level GA. Apply crossover operator and mutation operator to each upper level GA individual, and proceed to Step 7.

Step 7: If t_{uq} mod m_i (migration interval) = 0, after performing synchronization between the processors, apply migration. Return to Step 2 with $t_{uq} := t_{uq} + 1$.

4. Numerical Experiments

Numerical experiments are carried out in order to demonstrate the feasibility and the effectiveness of DGA. We apply DGA, PGA, and NGA to twelve types of two-level 0-1 programming problems as shown in Table 1. Each problem has five constraints.

Table 1
Problems used in the numerical experiments

Problem	DM1 variables	DM2 variables	Constraint strength
A	15	15	I (50%)
			II (70%)
			III (90%)
B	20	20	I (50%)
			II (70%)
			III (90%)
C	25	25	I (50%)
			II (70%)
			III (90%)
D	30	30	I (50%)
			II (70%)
			III (90%)

In this case, the elements A, B, $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{d}_1$, and $\mathbf{d}_2$ of the two-level 0-1 programming problem are selected at random from the closed interval $[10, 99]$, and the b_i element of $\mathbf{b}$ is set according to the equation

$$b_i = r_i\left(\sum_{j=1}^{n_1} a_{ij} + \sum_{k=1}^{n_2} b_{ik}\right), i = 1, \ldots, m. \quad (3)$$

Furthermore, a_{ij} represents the ij element of matrix A, and b_{ik} represents the ik element of matrix B. Here, r_i represents the strength of the constraint. In the strong constraint problem (I), a random number is determined at random from the closed interval $[0.45, 0.55]$, in the middle constraint problem (II), a random number is selected at random from the closed interval $[0.65, 0.75]$, and in the weak constraint problem (III), a random number is determined at random from the closed interval $[0.85, 0.95]$, respectively. The decimal portion of the b_i value is rounded off, and the result is stored as an integer value.

Next, the GA parameters are set for NGA, PGA, and DGA as follows. First, there are parameters that are used in common by all three methods, namely the population sizes of the upper level GA and the lower level GA, the crossover rate, the mutation rate, and the maximum number of generation, and those values are set as 120 for the population size, 0.9 for the crossover rate, 0.02 for the mutation rate, and 300 for the maximum number of generation. Next, for PGA and DGA, α is set to 0.25. The initial number of clusters, k, is set to 5, and d_{max} and d_{min}, used to measure the distance between individuals, are set to 2.5 and 1.0, respectively. In addition to these parameters, x_max and y_max are set to 100 and 5, respectively. The number of processors is set to 3. Finally, for DGA, we apply the random ring model to the communication topology and adopt Best-Hole model as the selection of emigrants and immigrants. The number of migration interval m_i is set to 15. The migration rate is set to 2.5%.

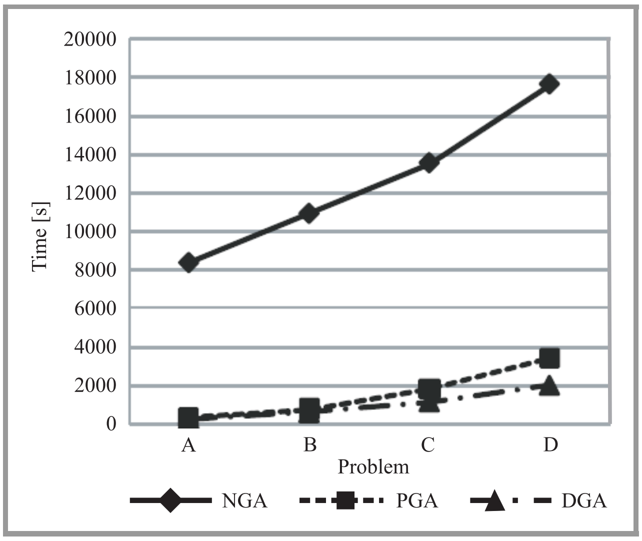


Fig. 3. Comparison of calculation times (the strong constraint problem (I)).

Table 2
Comparison of solution precisions

Problem	Constraint strength	DGA				PGA			
		best	worst	average	variance	best	worst	average	variance
A	I	1078	1078	1078.0	0.00	1078	1078	1078.0	0.00
	II	1377	1377	1377.0	0.00	1377	1377	1377.0	0.00
	III	1727	1727	1727.0	0.00	1727	1727	1727.0	0.00
B	I	1342	1342	1342.0	0.00	1342	1342	1342.0	0.00
	II	1764	1764	1764.0	0.00	1764	1764	1764.0	0.00
	III	2157	2157	2157.0	0.00	2157	2157	2157.0	0.00
C	I	1713	1713	1713.0	0.00	1713	1713	1713.0	0.00
	II	2221	2207	2214.6	43.84	2221	2207	2215.7	20.81
	III	2680	2680	2680.0	0.00	2680	2680	2680.0	0.00
D	I	1864	1857	1859.8	11.76	1864	1851	1856.8	12.36
	II	2480	2444	2469.6	253.44	2480	2439	2449.7	233.01
	III	2930	2930	2930.0	0.00	2930	2930	2930.0	0.00

Problem	Constraint strength	NGA				Enumeration
		best	worst	average	variance	
A	I	1078	1068	1077.0	9.00	1078
	II	1377	1377	1377.0	0.00	1377
	III	1727	1727	1727.0	0.00	1727
B	I	1342	1328	1340.4	17.24	-
	II	1764	1754	1763.0	9.00	-
	III	2157	2157	2157.0	0.00	-
C	I	1713	1713	1713.0	0.00	-
	II	2221	2207	2210.7	23.81	-
	III	2680	2672	2678.5	9.05	-
D	I	1864	1849	1856.9	11.29	-
	II	2480	2444	2470.0	80.80	-
	III	2930	2915	2927.0	20.00	-

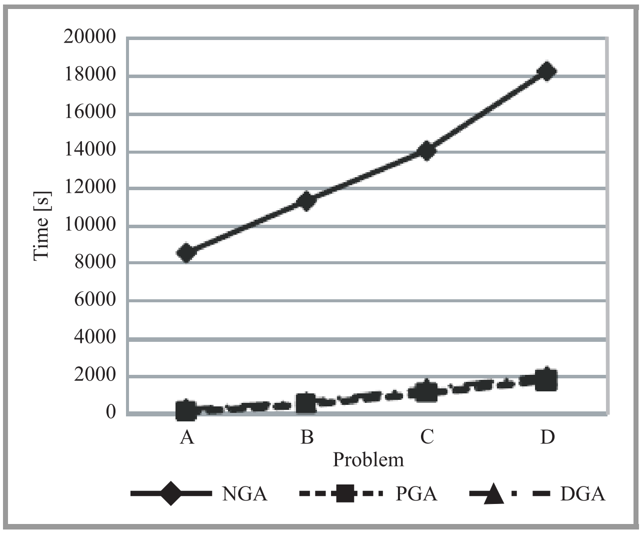


Fig. 4. Comparison of calculation times (the middle constraint problem (II)).

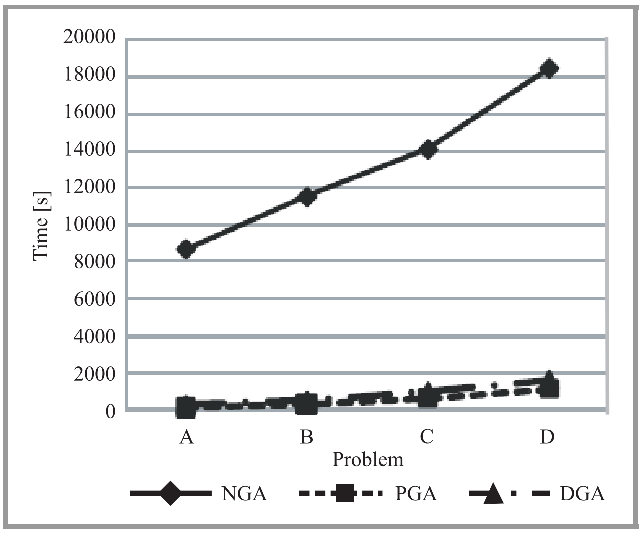


Fig. 5. Comparison of calculation times (the weak constraint problem (III)).

Next we will describe the experimental environment. The experiment is run on a personal computer with a 2.80 GHz CPU and running Windows XP. The compiler used is Microsoft Visual C++ 6.0.

For each problem, NGA, PGA, and DGA are run ten times each. The results are given in Table 2 through Table 5.

In Table 2, for all trials of Problem A, DGA, PGA, and NGA derive precise Stackelberg solutions. But, for the re-

Table 3
Comparison of calculation times

Problem	Constraint strength	DGA				PGA			
		best	worst	average	variance	best	worst	average	variance
A	I	220.78	351.25	269.82	1138.83	213.64	434.43	319.70	3611.94
	II	247.30	320.80	275.36	482.07	108.05	166.72	128.41	316.94
	III	215.46	281.27	253.94	427.14	62.20	99.46	82.38	191.60
B	I	394.09	801.84	599.33	12700.54	599.11	856.95	770.43	4550.59
	II	464.08	772.94	610.84	7149.91	419.83	569.50	504.90	2619.66
	III	345.83	633.73	473.91	5558.42	248.95	360.14	293.30	1186.14
C	I	975.14	1326.10	1149.18	10536.46	1728.47	1913.19	1824.04	4479.17
	II	1092.68	1569.18	1337.10	25015.18	1067.88	1199.30	1116.03	1888.83
	III	772.44	1191.86	965.45	16951.74	537.94	665.39	605.52	1484.93
D	I	1640.06	2426.27	2000.36	53181.01	3347.20	3547.55	3428.30	3115.52
	II	1568.27	2515.28	1980.26	81371.51	1753.85	1781.79	1767.90	97.18
	III	1012.63	2268.36	1619.37	89178.03	1073.29	1169.15	1125.58	848.64

Problem	Constraint strength	NGA				Enumeration
		best	worst	average	variance	
A	I	8379.75	8613.52	8414.59	4578.06	566.73
	II	8550.34	8668.06	8618.90	842.89	652.38
	III	8708.25	8714.92	8711.78	3.91	665.81
B	I	10882.38	11006.64	10954.14	1946.30	–
	II	11296.80	11467.94	11360.96	2888.93	–
	III	11447.75	11907.14	11520.65	17290.82	–
C	I	13467.95	13651.13	13573.35	1886.21	–
	II	14022.60	14067.24	14032.10	147.58	–
	III	14060.64	14090.92	14081.75	80.04	–
D	I	17532.97	17794.47	17646.13	4659.97	–
	II	18208.95	18263.34	18250.02	235.16	–
	III	18424.03	18541.39	18461.96	1912.68	–

Table 4
Lower level GA avoidance counts

Problem	Constraint strength	DGA				PGA			
		best	worst	average	variance	best	worst	average	variance
A	I	35558	34898	35283.7	46901.01	32008	27418	29742.4	1589311.44
	II	35539	35093	35291.8	24252.76	31094	26980	29362.6	1272588.24
	III	35634	35133	35371.4	24639.24	32006	29261	30431.2	948838.16
B	I	34983	33060	33995.6	265319.84	21852	14956	17355.0	3164457.60
	II	34757	32693	33959.8	317217.36	20845	14153	17304.8	4949162.76
	III	35274	34007	34667.5	121787.45	23258	17347	20721.5	3514280.45
C	I	32740	30118	31504.7	518636.01	8626	5577	6963.4	1213648.44
	II	32705	30017	31094.1	677375.69	5164	1599	3707.9	1823801.69
	III	33408	31591	32853.0	390069.20	16092	10170	13038.9	3552493.09
D	I	30696	26108	27780.8	1576362.56	1205	83	490.9	110786.89
	II	30819	26494	28875.6	2235743.84	847	114	292.5	46420.85
	III	32731	27108	30240.6	1688052.64	5939	3469	4376.3	514510.01

sults of Problem A-I, NGA is not possible to obtain precise Stackelberg solutions in some trials. Performance in deriving the best solutions to Problem B is equal for all three methods, but comparing the worst values and the average values shows that DGA and PGA present the best performance. Observe that the best values and the worst values obtained by DGA for Problem C and Problem D are equal or superior to the corresponding values obtained by NGA. Also, for Problem C and Problem D without Problem D-II, comparing the average values shows that DGA is superior to NGA. Finally, for Problem C and Problem D, DGA is superior or approximately equivalent to PGA in all results. When comparing the three methods with regards to the solution precision obtained, DGA is superior.

Table 5
Comparison of generations in which the best solution was found

Problem	Constraint strength	DGA				PGA			
		best	worst	average	variance	best	worst	average	variance
A	I	5	147	37.8	2210.96	3	42	13.9	130.09
	II	6	20	11.7	16.81	2	13	9.0	8.80
	III	5	20	11.0	20.80	3	16	9.3	18.81
B	I	5	158	54.4	1845.04	13	147	62.2	2204.76
	II	9	53	16.7	155.01	7	22	16.2	27.56
	III	9	29	16.6	35.24	9	30	18.9	35.69
C	I	15	71	23.0	258.60	15	59	38.3	180.01
	II	14	165	61.1	3067.89	50	292	141.1	6644.29
	III	13	247	75.6	6204.84	26	53	40.8	75.76
D	I	12	256	68.4	7015.44	62	291	157.3	5343.41
	II	23	299	117.7	9811.21	36	249	159.9	4262.09
	III	19	70	39.6	275.04	30	239	74.9	3441.69

Problem	Constraint strength	NGA				Enumeration
		best	worst	average	variance	
A	I	5	208	48.0	6130.80	–
	II	4	16	8.4	10.64	–
	III	3	126	22.4	1205.84	–
B	I	13	133	45.7	1728.41	–
	II	9	44	16.3	91.61	–
	III	9	63	19.5	223.25	–
C	I	11	35	20.5	34.85	–
	II	16	66	28.2	352.76	–
	III	13	26	18.8	17.36	–
D	I	19	108	34.7	666.61	–
	II	14	103	51.1	721.69	–
	III	23	284	64.1	5833.29	–

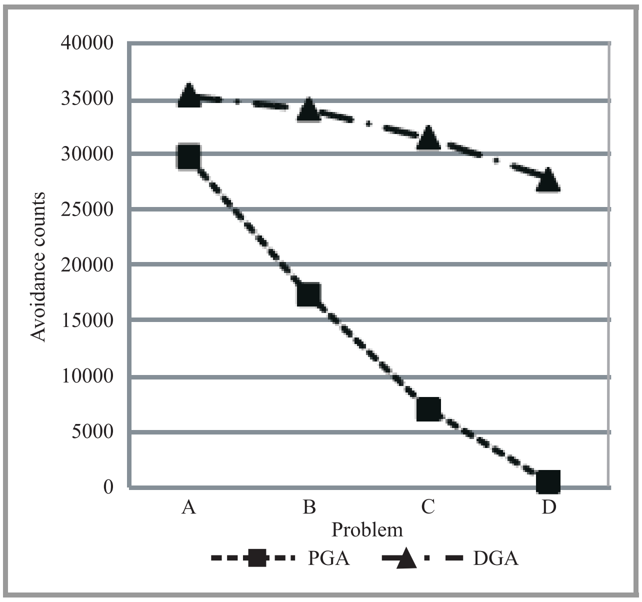


Fig. 6. Comparison of average avoidance counts (the strong constraint problem (I)).

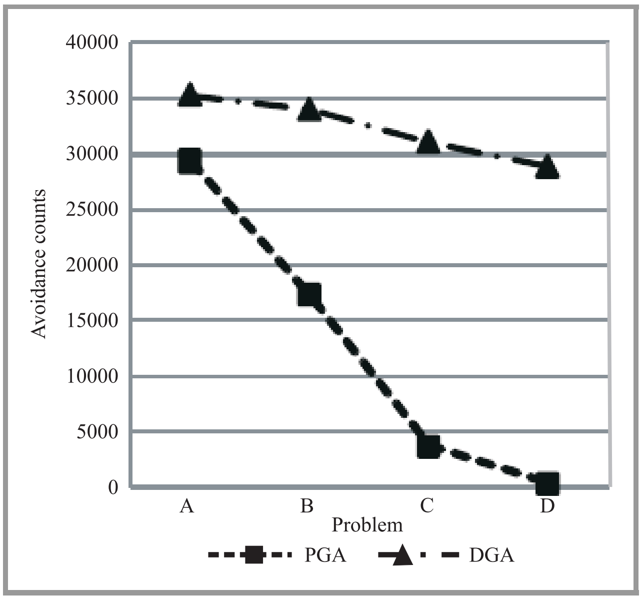


Fig. 7. Comparison of average avoidance counts (the middle constraint problem (II)).

Examining the calculation time results displayed in Table 3, for all trials of all problems, DGA and PGA are superior to NGA. Also, the calculation times of DGA are less than 15% of the calculation times of NGA.

Figure 3 shows the average calculation times for each size of the strong constraint problems (I). We can see that as compared to PGA, DGA is superior.

The average calculation times for each size of the middle constraint problems (II) are illustrated in Fig. 4. When comparing DGA and PGA, DGA is approximately equivalent or slightly inferior.

The average calculation times for each size of the weak constraint problems (III) are shown in Fig. 5. From Fig. 5, DGA is approximately equivalent or slightly inferior to PGA.

The results of the avoidance count for the lower level GA are shown in Table 4. From the results listed in this table, we see that with both DGA and PGA, as the scale of the problem increases the number of lower level GA avoidances is reduced.

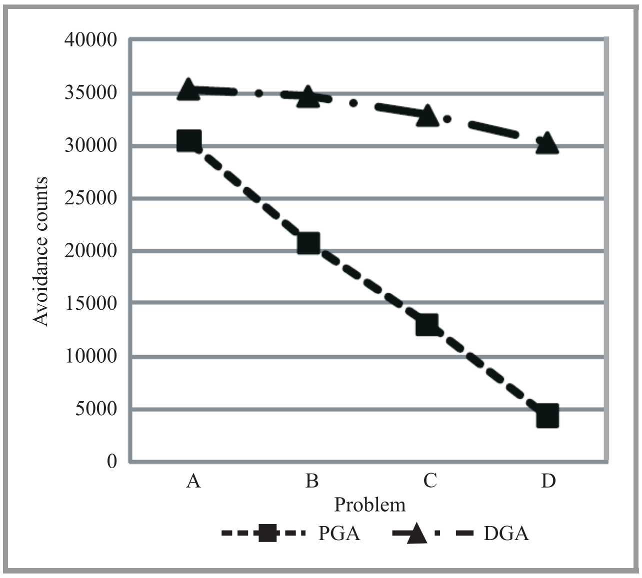


Fig. 8. Comparison of average avoidance counts (the weak constraint problem (III)).

The average avoidance counts for each size of the strong constraint problems (I), the middle constraint problems (II), and the weak constraint problems (III) are shown in Figs. 6–8. Examining the results displayed in these figures, we see that the avoidance counts of DGA gradually decrease as the size of the problem increases. On the other hand, the avoidance counts of PGA rapidly decrease.

Finally, the results of the comparison of generations in which the best solution was found are illustrated in Table 5. The results listed in Table 5 show that with both DGA and PGA, the generation deriving the best result is even later than with NGA, and so it is likely that by introducing cluster analysis methods the diversity of individuals within the population of the upper level GA would be maintained, and rapid population convergence avoided.

From the above results, compared to the other two computational methods, DGA is the superior computational method, both from a standpoint of solution precision and required calculation time.

5. Conclusion

This paper has focused on a two-level 0-1 programming problem in which there is not coordination between the decision maker at the upper level and the decision maker at the lower level. The authors have proposed a modified computational method that solves problems related to computational methods for obtaining the Stackelberg solution. Specifically, in order to improve the computational accuracy of approximate Stakelberg solutions and shorten the computational time of a computational method implementing GA proposed by the authors, a distributed genetic algorithm has been introduced with respect to the upper level GA, which handles decision variables for the upper level DM. Also, parallelization of the lower level GA has been performed along with parallelization of the upper level GA. The proposed algorithm has been improved in order to eliminate unnecessary computation during operation of the lower level GA, which handles decision variables for the lower level DM. In order to verify the effectiveness of the proposed method, numerical experiments have been carried out. From the results, we have shown that the proposed method is the superior to the other two computational methods, both from a standpoint of solution precision and required calculation time.

References

[1] K. Shimizu, Y. Ishizuka, and J. F. Bard, *Nondifferentiable and Two-Level Mathematical Programming*. Norwell: Kluwer, 1997.

[2] J. Bard and J.Moore, "An algorithm for the discrete bilevel programming problem", *Nav. Res. Log.*, vol. 39, pp. 419–435, 1992.

[3] J. Bard and J. Moore, "The mixed integer linear bilevel programming problem", *Oper. Res.*, vol. 38, pp. 911–921, 1990.

[4] W. P. Wen and Y. H. Yang, "Algorithms for solving the mixed integer two-level linear programming problem", *Comput. Oper. Res.*, vol. 17, pp. 133–142, 1990.

[5] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Norwell: Addison Wesley, 1989.

[6] M. Sakawa and M. Tanaka, *Genetic Algorithms*. Tokyo: Asakura Publishing, 1995.

[7] G. Anandalingam, R. Mathieu, C. L. Pittard, and N. Sinha, "Artificial intelligence based approaches for solving hierarchical optimization problems," in *Impacts of Recent Computer Advances on Operations Research,* R. Sharda, B. L. Golden, E. Wasil, O. Balci and W. Stewart, Eds. New York: Elsevier Science Publishing, 1989, pp. 289–301.

[8] I. Nishizaki and M. Sakawa, "Computational methods through genetic algorithms for obtaining Stackelberg solutions to two-level integer programming problems", *Cybernet. Syst. Int. J.*, vol. 36, pp. 565–579, 2005.

[9] I. Nishizaki and M. Sakawa, "Computational methods through genetic algorithms for obtaining Stackelberg solutions to two-level mixed zero-one programming problems", *Cybernet. Syst. Int. J.*, vol. 31, pp. 203–221, 2000.

[10] K. Niwa, I. Nishizaki, and M. Sakawa, "Decentralized two-level 0-1 programming through genetic algorithms with double strings", in *Proc. 2nd Int. Conf. Knowl. Int. Electron. Syst.*, 1998, vol. 2, pp. 278–284.

[11] K. Niwa, "Revised computational methods for using genetic algorithms for obtaining Stackelberg solutions to two-level 0-1 programming problems", in *Essays and Studies in Commemoration of the 40th Anniversary of the Founding of Hiroshima University of Economics*, Hiroshima University Economics, 2007, pp. 771–794 (in Japanese).

[12] K. Niwa, I. Nishizaki, and M. Sakawa, "Two-Level 0-1 Programming Using Genetic Algorithms and a Sharing Scheme Based on Cluster Analysis", in *Proc. Int. MultiConf. Eng. Comput. Sci.*, Hong Kong, China, 2008, pp. 1931–1936.

[13] K. Niwa, I. Nishizaki, and M. Sakawa, "Two-level 0-1 programming through parallel genetic algorithms", in *Proc. Int. MultiConf. Eng. Comput. Sci.*, Hong Kong, China, 2009, vol. 2, pp. 2000–2005.

[14] D. E. Goldberg and R. Lingle, "Alleles, loci, and the traveling salesman problem", in *Proc. First Int. Conf. Genet. Algorithms Appl.*, Hillsdale, USA, 1985, pp. 154—159.

[15] D. Abramson and J. Abela, "A parallel genetic algorithm for solving the school timetabling problem", in *Proc. 15th Austr. Comput. Sci. Conf.*, Hobart, Australia, 1992, vol. 14, pp. 1–11.

[16] E. Cantú-Paz, *Efficient and Accurate Parallel Genetic Algorithms.* Norwell, Massachusetts: Kluwer, 2000.

[17] T. C. Fogarty and R. Huang, "Implementing the genetic algorithm on transputer based parallel processing systems", in *Proc. First Int. Conf. Parallel Prob. Solv. Nat.*, Dortmund, Germany, 1990, pp. 145–149.

Keiichi Niwa received the Ph.D. degree in systems engineering from Hiroshima University in 2000. Since 2000, he has jointed Hiroshima University of Economics where he is currently a Associate Professor at Department for Information Systems in Business. His main research interests include meta-heuristics and integer programming.
e-mail: ki-niwa@hue.ac.jp
Department for Information Systems in Business
Faculty of Economics
Hiroshima University of Economics
Hiroshima, 731-0192, Japan

Tomohiro Hayashida is currently Assistant Professor at Department of Artificial Complex Systems Engineering, Graduate School of Engineering of the Hiroshima University, Japan. His current research interests are agent-based simulation analysis, game theory, and decision making theory.
e-mail: hayashida@hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering
Graduate School of Engineering
Hiroshima University
Higashi-Hiroshima, 739-8527, Japan

Masatoshi Sakawa received B.Sc., M.Sc., and Ph.D. degrees in applied mathematics and physics at Kyoto University, in 1970, 1972, and 1975, respectively. From 1975 he was with Kobe University where, since 1981, he was an Associate Professor in the Department of Systems Engineering. From l987 to 1990, he was a Professor of Department of Computer Science at Iwate University. At present he is a Professor of Hiroshima University, Japan, and is working with the Department of Artificial Complex Systems Engineering, Hiroshima University. His research and teaching activities are in the area of systems engineering, especially, mathematical optimization, multiobjective decision making, fuzzy mathematical programming and game theory.
e-mail: sakawa@hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering
Graduate School of Engineering
Hiroshima University
Higashi-Hiroshima, 739-8527, Japan

Paper

Coordination Games with Communication Costs in Network Environments

Ichiro Nishizaki, Tomohiro Hayashida, and Noriyuki Hara

Abstract—In this paper, we deal with a coordination game in a network where a player can choose both an action of the game and partners for playing the game. In particular, a player interacts with players connecting through a path consisting of multiple links as well as with players directly connecting by a single link. We represent decay or friction of payoffs with distance as communication costs, and examine the effect of the communication cost on behavior of players in the game and network formation. We investigate properties of equilibrium networks by classifying the link cost and the communication cost, and show diversity of the equilibrium networks.

Keywords—communication costs, coordination games, equilibrium, networks.

1. Introduction

Studies on formation of social systems and conventions have been accumulated by mathematically modeling social interaction between individuals through sequences of playing a game. In such game theoretic approaches, interaction between individuals is represented as playing coordination games, and selection of equilibria is considered in a dynamic process with perturbations or mutations. Recently, several articles have been devoted to similar attempts in network environments by allowing players to choose partners for playing the game as well as actions in the game [1], [2], [3], and our concern is also to consider this topic.

To examine formulation of conventions, Kandori *et al*. [4] and Young [5] deal with 2×2 coordination games which are repeatedly played by randomly matched pairs of players in a population. Ellison [6], Droste *et al*. [7], and Fagiolo [8] focus on locality of interaction between players. Oechssler [9], Ely [10], and Bhaskar and Vega-Redondo [11] consider location models where the population is divided into several groups and players choose which group to join. In the recent years, similar attempts in network environments have been attracting attention [1], [2], [3]. In such models, players are allowed to choose partners for playing the game as well as actions in the game. It should be noted that the network game models are related with studies on formation of networks [12], [13], [14].

Assuming that the link formation can be realized by an unilateral decision of a player and the player pays all the link cost, Goyal and Vega-Redondo [2] define Nash equilibrium networks, and consider stability of networks in the long run. For the case where interaction is restricted to a pair of two players connected by a direct link, they show the following result. For games where two players can obtain positive payoffs even in disequilibrium, the completely connected network is in equilibrium, and when the link cost is higher than the level of the payoff, networks with two completely connected components and the empty network are also in equilibrium. For the stability in the long run, if the link cost is smaller than a certain threshold, the risk dominant equilibrium networks are stochastically stable, and if the link cost is larger than it, the payoff dominant equilibrium networks are stochastically stable. Furthermore, they consider a network game model where any two players without a direct link are allowed to interact through two or more links connecting them. In this setting, they find that the unique stochastically stable structure of networks is a minimally connected network called a center-sponsored star network, and also find that there exists a certain threshold dividing two types of coordination: the risk dominant and the payoff dominant actions.

A study of Hojman and Szeidl [3] deals with a network game model with directed links similar to that of Goyal and Vega-Redondo [2]. In their model, it is assumed that interaction between two players connected not only by a single link but also by a path of multiple links is allowed, but a player can obtain payoffs only from interaction with other players to whom there are directed paths of links. They show that the structure of Nash equilibrium networks is wheel-shaped. For the long run stability, when the link cost is small and the disequilibrium payoff is positive, the risk dominant equilibrium is a unique stochastic stable network, and otherwise the payoff dominant equilibrium is uniquely stochastically stable on the condition that the payoff dominant equilibrium creates quite high positive gain or the degree of the risk dominance is small.

Although Goyal and Vega-Redondo [2] and Hojman and Szeidl [3] deal with interaction through a path, i.e., multiple consecutive links connecting players, it is assumed that the interaction between players connected by a path does not require any communication cost or such interaction is frictionless. However, it is natural to think that interaction with distant players costs and/or takes time much more than interaction between directly connected players. In this paper, we assume that a payoff arising from interaction through a path decreases with distance. This can be represented by a discount of the payoff or a communication cost of the network. While a discounted payoff is always nonnegative, the payoff from which the communication cost is subtracted may be negative. In this paper, employing a representation of decreasing the payoff by the communication cost, we deal with a network game model with interaction between distant players. Assuming that a link between two players is formed or maintained if the payoffs of both players do not decrease and they equally pay the link cost, we examine equilibrium networks. In Section 2, we introduce a network game model with the communication costs, and the equilibrium networks are shown in Section 3. Some concluding remarks are given in Section 4.

2. A Model

We deal with a network game model where a player chooses partners for interacting through direct links or paths consisting of multiple consecutive links. The interaction through a path requires the communication cost which increases with a distance between two players.

Let $N = \{1,\dots,n\}$ be the set of players which is called a population. The interaction between two players is represented by a 2×2 coordination game shown in Table 1. Let $a_i \in \{\alpha,\beta\}$ denote an action selected by player i. Each entry of the payoff table is a 2-dimensional vector and, the first element of the vector is a payoff of the row player and the second one is that of the column player. A vector $a = \{a_1,\dots,a_n\}$ of actions selected by all the players is called a profile of actions.

Table 1
Payoff table of a coordination game

Row player	Column player	
	α	β
α	(a,a)	(f,e)
β	(e,f)	(b,b)

Because we deal with a coordination game with conflict between the risk dominant equilibrium and the payoff dominant equilibrium, it is assumed that the following conditions are satisfied for the parameters of the payoffs given in Table 1.

$$a > b,\quad a > e,\quad b > f,\quad a+f < b+e. \qquad (1)$$

Thus, the outcome (α,α) is the payoff dominant equilibrium, and (β,β) is the risk dominant equilibrium.

The following notation is used. If there exists link ij between players i and j, $l_{ij} = 1$, and otherwise $l_{ij} = 0$. A set of links of player i is expressed by $l_i = (l_{i1},\dots,l_{i,i-1}, l_{i,i+1},\dots,l_{i,n}) \in \{0,1\}^{n-1}$. Because player i can decide to choose which links to maintain, the link set l_i of player i can be interpreted as a strategy for choices of links. A vector $l = (l_1,\dots,l_n)$ of link sets of all the players is a profile of links, and a set of links in the population is expressed by $g(l) = \{ij \mid l_{ij} = l_{ji} = 1\}$.

If $l_{ij} = 1$ or there exist a series of players $j_1,\dots,j_m$ such that $l_{ij_1} = \cdots = l_{j_k j_{k+1}} = \cdots = l_{j_m j} = 1$, it is said that there exists a path between players i and j. The existence of the path is denoted by $\bar{l}_{ij} = 1$, and the path is also expressed by $i \leftrightarrow j$. A set of paths of player i is expressed by $\bar{l}_i = (\bar{l}_{i1},\dots,\bar{l}_{i,i-1},\bar{l}_{i,i+1},\dots,\bar{l}_{in})$. For a path $i \leftrightarrow j$, the length of the path or the distance between players i and j is defined as the minimal number of links connecting i and j, and it is denoted by L_{ij}. A set of nodes corresponding to $g(l)$ is expressed by $N(g(l)) = \{i \mid \exists j, ij \in g(l)\}$. For a subset of links $g(l') \subset g(l)$, $g(l')$ is called a component of $g(l)$ if, for any pair of $i \in N(g(l'))$ and $j \in N(g(l'))$, there exists a path $i \leftrightarrow j$, and $ij \in g(l)$ implies $ij \in g(l')$.

To form or maintain link ij, players i and j need to pay the cost c. When there exits a path between players i and j, i.e., $\bar{l}_{ij} = 1$, they play the 2×2 coordination game paying the communication cost $d(L_{ij})$, where $d(\cdot)$ is a strictly increasing monotone function with the length of the path, and $d(1) = 0$. It is assumed that player i takes the same action a_i for all games with partners connected by direct links or paths. Let $a_{-i} = (a_1,\dots,a_{i-1},a_{i+1},\dots,a_n)$ and $l_{-i} = (l_1,\dots,l_{i-1},l_{i+1},\dots,l_n)$ be, respectively, a profile of actions and a profile of links for $N_{-i} = N\setminus\{i\}$ which is the set of players except for player i. Then, a utility of player i with a strategy (a_i,l_i) is written by

$$\pi_i((a_i,l_i),(a_{-i},l_{-i})) = \sum_{j:\bar{l}_{ij}=1} u(a_i,a_j) - \sum_{j:l_{ij}=1} c - \sum_{j:\bar{l}_{ij}=1} d(L_{ij}), \quad (2)$$

where $u(a_i,a_j)$ is player i's payoff of the game shown in Table 1 when player i chooses action $a_i \in \{\alpha,\beta\}$ and player j chooses action $a_j \in \{\alpha,\beta\}$. For concise representation, the part of costs in (2) is defined by

$$D_i = \sum_{j:l_{ij}=1} c + \sum_{j:\bar{l}_{ij}=1} d(L_{ij}). \qquad (3)$$

Let $A \triangleq \{\alpha,\beta\}$ and $L \triangleq \{0,1\}^{n-1}$ denote the strategy sets of actions and links, respectively. The strategy set of a player is represented by $X \triangleq A\times L$. A strategy of player i is represented by a pair of an action and a set of links, and it is denoted by $s_i = (a_i,l_i) \in X$, $a_i \in A$, $l_i \in L$. A strategy profile $s = (s_1,s_2,\dots,s_n) \in X^n$ of all the players indicates a state of the network, and we call a strategy profile s a network or a state. We assume that each of the players knows a state s and can calculate the utilities of the other players $\pi_j(s)$, $j \neq i$ as well as the utility $\pi_i(s)$ of self. Moreover, we assume that link ij is formed only if both of the utilities of i and j do not decrease by forming link ij. Namely, when player i is selected to revise a strategy, for any player j such that $l_{ij} = l_{ji} = 0$ at the state s before revising the strategy of i and $l_{ij} = l_{ji} = 1$ at the state s' after revising it, the condition $\pi_j(s') \geq \pi_j(s)$ must be satisfied. It is said that the strategy s'_i of player i is feasible if this condition is satisfied, and a set of feasible strategies of player i is denoted by $X_i \subset X$. It follows that player i chooses a strategy s'_i among the feasible strategy set X_i, i.e., $s'_i \in X_i$. Because we assume that each player always chooses a strategy among the feasible strategy set, Nash equilibrium networks can be defined as follows.

Definition 1: A state $\hat{s} = (\hat{s}_1, \ldots, \hat{s}_n)$ is said to be an equilibrium state or an equilibrium network if, for any player $i \in N$ and any feasible strategy $s_i \in X_i$, the following condition holds:

$$\pi_i(\hat{s}_i, \hat{s}_{-i}) \geq \pi_i(s_i, \hat{s}_{-i}). \quad (4)$$

Such a strategy $\hat{s}_i = (\hat{a}_i, \hat{l}_i)$ is called an equilibrium strategy of player *i*, and a component in an equilibrium network is called an equilibrium component.

In this paper, focusing on actions of players and the formation of links, we examine static equilibrium networks in the network game with the communication cost. At the beginning of the examination, we give some definitions on actions of players and structures of networks. We call a player who chooses action α an α-player, and a player who chooses action β a β-player. A player holding only one link is called a leaf player, and a player holding no link is called an isolated player. If all of players who belong to a component $g(l')$ choose action α, the component $g(l')$ is called an α-component, and similarly a β-component is defined. If any pairs of players $i, j \in N(g(l'))$ are connected by a direct link, i.e., $l'_{ij} = l'_{ji} = 1$, the component $g(l')$ is said to be completely connected. If a component is divided by severing any link in the component, the component is said to be minimally connected. If after severing a ceratin link in a component, the component is not divided, there should exist a loop of links in the component. A component without any leaf player is called a leafless component. In particular, if all the players in a component have only two links and they are arranged like a circle, the component is called a ring component. These definitions are given for components, and similar definitions are also given for networks.

3. Equilibrium Networks

We deal with coordination games in network environments and examine equilibria of networks in this section. If an equilibrium network is not the empty network, there exists at least one component. Then, we first characterize equilibrium components.

Lemma 1: In any equilibrium component, all the players in the component choose the same action.

The proof of Lemma 1 is given in Appendix, and the proofs of the subsequent lemmata and theorem are also given in Appendix. While we focused on behavior of players in Lemma 1, the next lemma deals with structures of equilibrium components. Let C and $|C|$ denote a component and the number of players in the component, respectively. As shown in Lemma 1, all players choose the same action in an equilibrium component, and let the action be $x \in A = \{\alpha, \beta\}$. First, we consider the case where the link cost is smaller than the payoff obtained by coordination of choices, i.e., $c < u(x,x)$. Let L denote a distance between players i and j, and let k denote any player on path $i \leftrightarrow j$. Because a distance between players i and k decreases or does not change when link ij is formed, the sum of the communication costs over path $i \leftrightarrow j$ decreases, and the reduced cost is calculated as follows:

$$RC(L) = \sum_{k=\lceil L/2 \rceil + 1}^{L} d(k) - \sum_{k=1}^{\lfloor L/2 \rfloor} d(k), \quad (5)$$

where $\lceil x \rceil$ and $\lfloor x \rfloor$ mean the minimal integer larger than or equal to x and the maximal integer smaller than or equal to x, respectively.

Lemma 2: For a given action $x \in A = \{\alpha, \beta\}$, if $c < u(x,x)$, an equilibrium x-component C has the following structures.

(1) If $c < d(2)$, an equilibrium component is completely connected, and vice versa.

(2) If $d(2) \leq c$ and $c \leq RC(|C| - 1)$, there exists an equilibrium component which is not completely connected. The maximal length of a path in the equilibrium component is the largest number L satisfying $RC(L) \leq c$ and $L < |C|$.

(3) If $d(2) \leq c$ and $RC(|C| - 1) < c$, an equilibrium component is minimally connected.

From $b < a$, $c < u(\beta,\beta)$ implies $c < u(\alpha,\alpha)$. Then, the result of Lemma 2 is valid for α- and β-components if $c < b$, it is valid for α-components if $b < c < a$, and it is not valid for either of them if $a < c$.

For (3) of Lemma 2, any minimally connected component is not always an equilibrium component. As a counterexample, consider a component where 4 players are in line. From $|C| = 4$ and the condition of (3) of Lemma 2, the inequality $c > d(3) - d(2)$ holds. The utility of player i at the end of the line is $\pi_i = 3u(x,x) - c - d(2) - d(3)$, and from $c > d(3) - d(2)$, it satisfies the following inequality.

$$\begin{aligned}\pi_i &= 3u(x,x) - c - d(2) - d(3) \\ &< 3u(x,x) - (d(3) - d(2)) - d(2) - d(3) \\ &= 3u(x,x) - 2d(3).\end{aligned}$$

Then, the utility π_i of player i is negative when $3u(x,x) - 2d(3) < 0$. For example, when $c = 4.5, u(x,x) = 5$, $d(2) = 4, d(3) = 8$, because the condition of Lemma 2: $d(2) \leq c < u(x,x)$ and $d(3) - d(2) < c$ and the above condition: $3u(x,x) - 2d(3) < 0$ are satisfied, the utility π_i of player i is negative. Thus, the best response of player i is to sever the link, and it follows that the component is not an equilibrium component.

We also show an example of a minimally connected component which is an equilibrium component. Consider a star-shaped minimally connected component where player 1 is the center and players 2, 3, and 4 are peripheries. Because a distance between any pair of players is at most two and $d(2) \leq c$, a new link is never formed. From $c < u(x,x)$, the utility of player 1, $\pi_1 = 3(u(x,x) - c)$, is positive, and those of the other players, $\pi_2 = \pi_3 = \pi_4 = 3u(x,x) - c - 2d(2)$, are positive. When any link is severed, the utility of player 1 decreases by $u(x,x) - c$. At this moment, the counterpart is isolated and her utility becomes zero. Thus, because in the star-shaped component, a new link is not formed and any of the existing links is not severed, it is an equilibrium component.

In Lemma 2, we have considered the structure of components, and in the next lemma, we examine whether or not there exist multiple components in an equilibrium network.

Lemma 3: For a given action $x \in A = \{\alpha, \beta\}$, if $c < u(x,x)$, in an equilibrium network, there exists only one x-component, and any x-player is connected with some other player.

Next, we consider the case where the link cost is larger than the payoff obtained by coordination of choices, i.e., $c > u(x,x)$. In general, an x-component is not likely to be formed in this case, but it could be maintained because players can obtain a positive payoff arising from interaction between distant players through paths if the communication cost is relatively small.

Lemma 4: For a given action $x \in A = \{\alpha, \beta\}$, if $c > u(x,x)$, in an equilibrium network, the necessary condition for existing one or more x-components is that the condition $d(2) \le c \le RC(2l)$ is satisfied, and there exists an integer $l \in (1, n/2]$ satisfying the condition

$$c - u(x,x) \le (l-1)u(x,x) - \sum_{k=1}^{l} d(k). \qquad (6)$$

Moreover, such equilibrium x-components are leafless.

From the above discussion, we have the following results.

Theorem 1: Assuming that the communication cost $d(\cdot)$ is a strictly increasing monotone function with a distance between any pair of players, we can characterize equilibrium networks as follows.

(1) In the case of $c < b$:

(a) If $c < d(2)$, an equilibrium network is the completely connected α-network or the completely connected β-network.

(b) If $d(2) \le c$ and $c \le RC(n-1)$, an equilibrium network is an incompletely connected α-network or an incompletely connected β-network. The maximal length of a path in the equilibrium network is the largest number L satisfying the conditions $RC(L) \le c$ and $L < |C|$.

(c) If $d(2) \le c$ and $c > RC(n-1)$, an equilibrium network is a minimally connected α-network or a minimally connected β-network.

(2) In the case of $b < c < a$:

(a) If $c < d(2)$, an equilibrium network is a completely connected α-network or an empty β-network.

(b) If $d(2) \le c$ and $c \le RC(n-1)$, an equilibrium network is an incompletely connected α-network or an empty β-network. Moreover, if there exists an integer $l \in (1, n/2]$ satisfying the condition $c - b \le (l-1)b - \sum_{k=1}^{l} d(k)$, a network with one or more leafless β-components can be an equilibrium network, and it may include one α-component. The maximal length of a path in the equilibrium network is the largest number L satisfying the conditions $RC(L) \le c$ and $L < |C|$.

(c) If $d(2) \le c$ and $c > RC(n-1)$, an equilibrium network is a minimally connected α-network or an empty β-network.

(3) In the case of $a < c$:

(a) If $c < d(2)$, an equilibrium network is an empty network.

(b) If $d(2) \le c$ and $c \le RC(n-1)$, an equilibrium network is an empty network. Moreover, if there exists an integer $l_a \in (1, n/2]$ satisfying the condition $c - a \le (l_a - 1)a - \sum_{k=1}^{l_a} d(k)$, a network with one or more leafless α-components can be an equilibrium network; and if there exists an integer $l_b \in (1, n/2]$ satisfying the condition $c - b \le (l_b - 1)b - \sum_{k=1}^{l_b} d(k)$, a network with one or more leafless β-components can be an equilibrium network.

(c) If $d(2) \le c$ and $c > RC(n-1)$, an equilibrium network is an empty network.

The structures of equilibrium networks shown in Theorem 1 are summarized in Table 2. In Theorem 1, the structures of equilibrium networks are characterized by the relation between the link cost c and the communication cost $d(2)$ of distance 2. In general, when the link cost c is smaller than the payoff of the game such as the payoff dominant equilibrium payoff a and the risk dominant equilibrium payoff b, a link is formed. Moreover, when the communication cost $d(\cdot)$ is large, compared to the link cost c, a link between any pair of players is likely to be formed, and when the communication cost $d(\cdot)$ is small, the number of links decreases because not direct links but paths allow players to interact with other players at lower costs. Moreover, comparing the link cost c and the communication cost $d(2)$ of distance 2 reveals whether or not players should form a link. By analysis taking into account the communication cost of distances larger than 2, we can examine the length of a path in a network.

In (1) of Theorem 1, because the link cost c is smaller than the payoff b of the risk dominant equilibrium, an equilibrium network is either an α-network which means the risk dominant equilibrium or a β-network which means the payoff dominant equilibrium. For the case of (a), because interaction by using direct links is profitable due to $c < d(2)$, the completely connected network is formed. In the case of (c), interaction by using direct links is costly

Table 2
Equilibrium networks

Case	Condition	(1) $c < b$	(2) $b < c < a$	(3) $a < c$
(a)	$c < d(2)$	complete α complete β	complete α empty β	empty
(b)	$d(2) \leq c$ $c \leq RC(n-1)$	incomplete α incomplete β	incomplete α empty β (leafless β)	empty (leafless α) (leafless β)
(c)	$d(2) \leq c$ $c > RC(n-1)$	minimal α minimal β	minimal α empty β	empty
(leafless α) or (leafless β) is conditional.				

compared to interaction through paths because of $d(2) \leq c$, and from $c > RC(n-1)$, any loop is not formed. Thus, an equilibrium network is minimally connected.

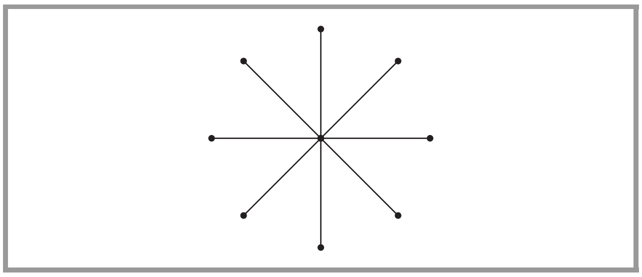


Fig. 1. A minimally connected (star) network.

As an example of a minimally connected network, a star-shaped network is given in Fig. 1. In this network, a new link between players of two units of distance is not formed because forming such links results in decrease of the utilities of the players. Moreover, any existing link is not severed because by severing a link the utilities of two players decrease by at least $b-c$. Thus, this type of networks are in equilibrium.

In the case of (b), because of $d(2) \leq c$ and $c \leq RC(n-1)$, an equilibrium network is not completely connected but it includes a loop. The density of networks, i.e., the number of links depends on the relation between the link cost c and the communication cost $d(\cdot)$, and it can be characterized by the maximal length of paths.

An example of an incompletely connected equilibrium network is given in Fig. 2. The length of a path in this network is at most 3, and if $\max\{d(2), RC(3)\} \leq c$, a new link is not formed. Because severing any link in the loop makes a pair of players of 4 units of distance, the utilities of them decrease by severing the link if $c < RC(4)$. For example, let $d(k) = k$. Then, from $RC(3) = 1$ and $RC(4) = 5$, this network is an equilibrium network if $2 \leq c < 5$. On the other hand, the condition of Theorem 1, $d(2) \leq c \leq RC(8)$, can be written as $2 \leq c \leq 16$. Then, when $c = 2, 3, 4$, both of the equilibrium condition of this network, $\max\{2, RC(3)\} \leq c < RC(4)$, and the condition of Theorem 1, $d(2) \leq c \leq RC(8)$, are satisfied simultaneously. With the above mentioned parameters, the incompletely connected network given in Fig. 2 is an equilibrium network.

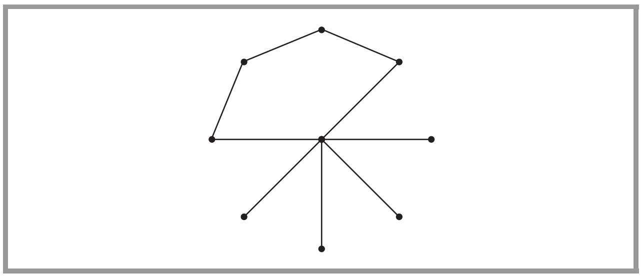


Fig. 2. An incomplete network.

In (2) of Theorem 1, because the link cost c is larger than the payoff b of the risk dominant equilibrium and it is smaller than the payoff a of the payoff dominant equilibrium, it is supposed that an equilibrium network is an α-network but a β-network is not the case. However, it is shown that if the condition given in the theorem is satisfied, some β-network can be an equilibrium. For the case of (a), from $c < d(2)$, the completely connected network is formed when all the players choose action α, and as a special case, a state where all the players choose action β in the empty network is an equilibrium network. In the case of (c), because interaction by using direct links is costly compared to interaction through paths for the same reason as in (1), a minimally connected equilibrium α-network is formed. The empty β-network is also an equilibrium. In the case of (b), because of $d(2) \leq c$ and $c \leq RC(n-1)$, an equilibrium network is not completely connected. Moreover, as special structures of equilibrium networks, besides the empty β-network, networks with one or more leafless β-components and networks with one α-component and one or more leafless β-components can be equilibria. However, an equilibrium network with a β-component can exist only when the condition given in the theorem is satisfied. Even if the link cost c is larger than the payoff b of the game, the utilities of players may become positive because of interaction between distant players, and then leafless β-components can be included in an equilibrium network. It is noted that there does not exist a leaf player in a β-component because the utility of a player who shares a link with the leaf player increases by severing the link.

We give some examples of equilibrium networks with multiple components in Figs. 3 and 4. In a network shown in Fig. 3, assume that the payoffs of the game are set at $a = 11$, $b = 6.71$, $e = 1$, and $f = 0$, and the costs of the link and the communication are set at $c = 10$, $d(2) = 5$, $d(3) = 5.1$, and $d(4) = 20$, respectively.

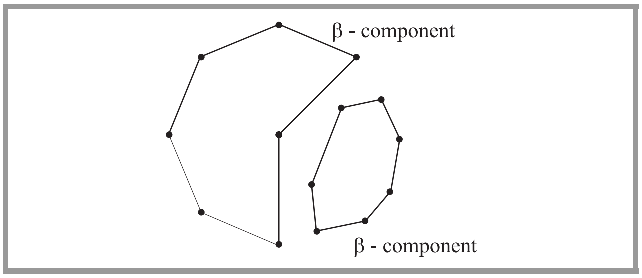


Fig. 3. A network with two β-components.

In Fig. 3, a β-network with two β-components is given. Forming a link within a component results in decrease of the utility, and therefore a new link within a component is not formed. Because by severing a link the distance between players becomes more than or equal to 4 and then the cost of a player increases, any link is not severed. Next, consider formation of a link between components. Because the utilities of players decrease by forming a link connecting two players in different components, a link between two components is not formed. In the case where links within a component and between components are formed simultaneously, no player is better off. Moreover, by moving to the other component, the utility of any player does not increase. Thus, the network with two β-components in Fig. 3 is an equilibrium network.

A network with coexistence of an α-component and a β-component is shown in Fig. 4. Let the payoffs of the game be $a = 8$, $b = 6$, $e = 1$, and $f = 0$, and suppose that the costs of the link and the communication are $c = 7$, $d(2) = 3$, $d(3) = 4$, and $d(4) = 12$. With these parameter values, by a similar consideration, one finds that the network with coexistence of an α-component and a β-components in Fig. 4 is an equilibrium network.

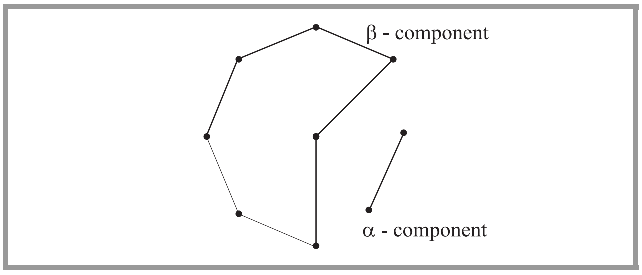


Fig. 4. A network with α- and β-components.

In (3) of Theorem 1, because of $c > a$, an equilibrium network is generally an empty network where the actions of players are unspecified. However, if $d(2) \leq c \leq RC(n-1)$ is satisfied, a leafless component can be included in an equilibrium network. In the case of (a), from $c < d(2)$, one finds that there does not exist an integer l satisfying the condition (6) in Lemma 4 because the left hand side of (6) is negative, and then only the empty networks are equilibria. For the case of (c), similarly, the condition of Lemma 4 is not satisfied due to $c > RC(n-1)$, and then there does not exist any nonempty equilibrium network. In contrast, in the case of (b), from $d(2) \leq c \leq RC(n-1)$, if players can obtain larger payoffs from the interaction between distant players through paths, besides the empty equilibrium networks, there can exist an equilibrium network with a leafless component.

4. Conclusions

In this paper, we dealt with the network game model in which a player can choose partners for playing the underlying coordination games as well as an action of the game, assuming that interaction between distant players is possible, but it requires the payment of the communication cost. We examined influence of the communication cost on the behavior of players in the game and the structure of networks. We showed a diversity of the equilibrium networks. The relevant studies [1], [2], [3] examine the long-run stability of the equilibrium networks. Naturally, it is interesting to investigate the stability of the equilibrium networks in the network game model considered in this paper, and we will intend to deal with this topic.

Appendix

Proof of Lemma 1

Assume that a component with α-players and β-players is an equilibrium component, and let n_α and n_β denote the numbers of α-players and β-players, respectively. Then, the utility of α-player i defined by (2) is expressed by

$$\pi_i^\alpha = \left\{(n_\alpha - 1)u(\alpha,\alpha) + n_\beta u(\alpha,\beta)\right\} - D_i\,, \qquad (7)$$

where D_i is the total cost defined by (3). Assume that player i changes his action of the game from α to β, and let π_i^β denote the utility of player i at the time. Because the structure of the component is the same as before, D_i does not change, and then the utility π_i^β is expressed by

$$\pi_i^\beta = \left\{(n_\alpha - 1)u(\beta,\alpha) + n_\beta u(\beta,\beta)\right\} - D_i\,. \qquad (8)$$

Because the component is in equilibrium before player i changes his action, the inequality $\pi_i^\alpha \geq \pi_i^\beta$ holds, and from Eqs. (7) and (8), one finds that

$$(n_\alpha - 1)a + n_\beta f \geq (n_\alpha - 1)e + n_\beta b\,. \qquad (9)$$

Similarly, for β-player j in the component, the following inequality holds.

$$n_\alpha e + (n_\beta - 1)b \geq n_\alpha a + (n_\beta - 1)f\,. \qquad (10)$$

From Ineqs. (9) and (10), one finds that

$$-(b-f) \geq a-e\,. \tag{11}$$

Because the inequality (11) is inconsistent with the assumption (1) of the payoffs of the game, $a>e$ and $b>f$, the component with α-players and β-players is not an equilibrium component. ■

Proof of Lemma 2

Let L be the distance between players i and j, i.e, $L_{ij}=L$, $2 \leq L < |C|$. First, we show the condition that link ij is not formed. The total cost of player i for path $i \leftrightarrow j$ is calculated as follows:

$$c+\sum_{k=1}^{L} d(k)\,. \tag{12}$$

If link ij is formed, i.e., $l_{ij}=1$, the total cost of player i changes to

$$c+\sum_{k=1}^{\lceil L/2\rceil} d(k)+\sum_{k=1}^{\lfloor L/2\rfloor} d(k)+c\,. \tag{13}$$

Thus, if the cost (13) after link ij is formed is larger than the original cost (12), i.e.,

$$c \geq \sum_{k=\lceil L/2\rceil+1}^{L} d(k)-\sum_{k=1}^{\lfloor L/2\rfloor} d(k)\,, \tag{14}$$

then link ij is not formed.

To prove (1), assume that an equilibrium component is not completely connected. Then, there exists at least one pair of players i and j such that $L_{ij} \geq 2$. In this case, because link ij is not formed, the condition (14) is satisfied. When $L=2$, one finds that $c \geq d(2)-d(1)=d(2)$, which is inconsistent with the assumption of (i): $c<d(2)$. Thus, if there exists an equilibrium network, its component should be completely connected.

Consider a completely connected component. The utility of player i arising from the interaction with player j is $u(x,x)-c-d(1)$. Then, if link ij is severed, the utility of player i changes to $u(x,x)-d(2)$. Because, from $d(1)=0$ and $c<d(2)$, severing link ij results in decrease of the utility, the completely connected component is an equilibrium component.

Consider the case of (3) before the case of (2), and assume that there exists an equilibrium component with a loop consisting of L players. Let i and j be adjacent players in the loop. Because of the assumption of equilibrium, severing link ij results in increase of the cost, and one finds that the following inequality holds:

$$2c+\sum_{k=1}^{\lceil L/2\rceil} d(k)+\sum_{k=1}^{\lfloor L/2\rfloor} d(k) \leq c+\sum_{k=1}^{L} d(k)\,.$$

Namely, we have

$$c \leq \sum_{k=\lceil L/2\rceil+1}^{L} d(k)-\sum_{k=1}^{\lfloor L/2\rfloor} d(k)=RC(L)\,. \tag{15}$$

Because $d(\cdot)$ is a strictly monotone increasing function and the right hand side of (15) is at most $RC(|C|-1)$, for any $L<|C|$, the inequality (15) is inconsistent with the assumption of (3): $c>RC(|C|-1)$. Thus, there does not exist any loop in an equilibrium component, and then it is a minimally connected component.

Finally, for the case of (2), assume that an equilibrium component is completely connected. Because severing a link results in increase of the cost, it follows that $2c<c+d(2)$. This is inconsistent with the assumption of (2): $d(2) \leq c$, and therefore an equilibrium component is not completely connected. To show that an equilibrium component is not restricted to be minimally connected, we demonstrate that a ring shaped component can be an equilibrium component.

Let C and $|C|$ denote a ring shaped component and the number of players in the component, respectively. First, we give the condition that severing a link of player i results in increase of the cost of player i. The total cost D_i of player i in the component is

$$D_i=c+\sum_{k=1}^{\lceil(|C|-1)/2\rceil} d(k)+\sum_{k=1}^{\lfloor(|C|-1)/2\rfloor} d(k)+c\,. \tag{16}$$

When player i severs a link, the total cost of player i changes to

$$D_i^-=c+\sum_{k=1}^{(|C|-1)} d(k)\,. \tag{17}$$

Thus, the condition that severing a link of player i results in increase of the cost of player i is $D_i<D_i^-$, i.e.,

$$c<\sum_{k=\lceil(|C|-1)/2\rceil+1}^{|C|-1} d(k)-\sum_{k=1}^{\lfloor(|C|-1)/2\rfloor} d(k)\,. \tag{18}$$

Second, consider the condition that forming a new link results in increase of the cost. Let the distance between players i and j be $L_{ij}=L$. When player i is directly connected with player j, two loops are formed; one loop has $L+1$ players, and the other has $|C|-L+1$ players. At this time, the cost of player i changes to

$$\begin{aligned} D_i^+=c&+\sum_{k=1}^{\lceil L/2\rceil} d(k)+\sum_{k=1}^{\lfloor L/2\rfloor} d(k)+c\\ &+\sum_{k=1}^{\lceil(|C|-L)/2\rceil} d(k)+\sum_{k=1}^{\lfloor(|C|-L)/2\rfloor} d(k)+c\,. \end{aligned} \tag{19}$$

Thus, the condition that forming a new link of player i results in increase of the cost of player i is $D_i<D_i^+$, i.e.,

$$\begin{aligned} c>&\sum_{k=\lceil L/2\rceil+1}^{\lceil(|C|-1)/2\rceil} d(k)-\sum_{k=1}^{\lfloor L/2\rfloor} d(k)\\ &+\sum_{k=\lceil(|C|-L)/2\rceil+1}^{\lfloor(|C|-1)/2\rfloor} d(k)-\sum_{k=1}^{\lfloor(|C|-L)/2\rfloor} d(k)\,. \end{aligned} \tag{20}$$

Third, suppose that player i severs link ij and forms a new link is, $s \neq j$, where $L_{is} = L$. At this time, the cost of player i changes to

$$D_i^{-+} = \sum_{k=1}^{L} d(k) + c + \sum_{k=1}^{\lceil (|C|-L)/2 \rceil} d(k) + \sum_{k=1}^{\lfloor (|C|-L)/2 \rfloor} d(k) + c\,. \quad (21)$$

Because $d(\cdot)$ is a strictly monotone increasing function, one finds that $D_i^{-+} < D_i$, namely, the cost of player i decreases by this operation. As for the cost of player s, it is

$$D_s^{-+} = c + \sum_{k=1}^{L} d(k) + c + \sum_{k=1}^{\lceil (|C|-L)/2 \rceil} d(k) + \sum_{k=1}^{\lfloor (|C|-L)/2 \rfloor} d(k) + c\,, \quad (22)$$

and therefore $D_s^{+} < D_s^{-+}$. Because if (20) is satisfied, $D_s < D_s^{+} < D_s^{-+}$ holds, player s rejects player i's offer to form link is. Therefore, if both Ineqs. (18) and (20) are satisfied simultaneously, the ring shaped component C is an equilibrium component. Moreover, because from the proof of (1), if $c < d(2)$, an equilibrium component is completely connected, $d(2) \leq c$ should be also satisfied. Thus, under the following condition, a component which includes a loop but is not completely connected can be an equilibrium component.

$$\max\left\{ d(2), \sum_{k=\lceil L/2 \rceil+1}^{\lceil (|C|-1)/2 \rceil} d(k) - \sum_{k=1}^{\lfloor L/2 \rfloor} d(k) + \sum_{k=\lceil (|C|-L)/2 \rceil+1}^{\lfloor (|C|-1)/2 \rfloor} d(k) - \sum_{k=1}^{\lfloor (|C|-L)/2 \rfloor} d(k) \right\} < c < \sum_{k=\lceil (|C|-1)/2 \rceil+1}^{|C|-1} d(k) - \sum_{k=1}^{\lfloor (|C|-1)/2 \rfloor} d(k) = RC(|C|-1)\,. \quad (23)$$

As for the maximal length of a path, because (14) is the condition that link ij is not formed when the distance between players i and j is L, the maximal length is derived straightforwardly from the condition. ■

Proof of Lemma 3

Assume that a network with two or more x-components is in equilibrium, and select any two components C_1 and C_2 in them. Let $|C_1|$ and $|C_2|$ be the numbers of players in C_1 and C_2, respectively. Consider player i in C_1 who selects action a_i and player j in C_2 who selects action a_j. If player i severs all his links in C_1, takes the same action as a_j, and offers to form a new link with player j in C_2, then player j accepts the offer because of $u(a_j, a_j) > c$, and therefore any player can move another component.

For the case of $c < d(2)$, because any component is completely connected, the condition that player i does not have any incentive to move from C_1 to C_2 is as follows.

$$(|C_1|-1)(u(a_i,a_i)-c) \geq |C_2|(u(a_j,a_j)-c)\,. \quad (24)$$

For player j, the following similar condition is obtained.

$$(|C_2|-1)(u(a_j,a_j)-c) \geq |C_1|(u(a_i,a_i)-c)\,. \quad (25)$$

Thus, if the following inequality condition is satisfied, both players do not move.

$$0 \geq (u(a_i,a_i)-c) + (u(a_j,a_j)-c)$$

However, because $\min\{u(a_i,a_i), u(a_j,a_j)\} > c$, the above inequality does not hold, and therefore in the case of $c < d(2)$ a network with two or more x-components is not an equilibrium network.

For the case of $c \geq d(2)$, assume that player i forms a new link with player k in C_2 who has a link with player j. At this time, the utility of player i is represented by

$$\pi_i = |C_2| u(a_j,a_j) - (D_j + d(2))\,,$$

where D_j denotes the total cost of player j. If the following inequality condition is satisfied, player i does not have any incentive to move from C_1 to C_2.

$$(|C_1|-1)u(a_i,a_i) - D_i \geq |C_2| u(a_j,a_j) - (D_j + d(2))\,. \quad (26)$$

For player j, the following similar condition is obtained.

$$(|C_2|-1)u(a_j,a_j) - D_j \geq |C_1| u(a_i,a_i) - (D_i + d(2))\,. \quad (27)$$

Thus, if the following inequality condition is satisfied, both players do not move.

$$0 \geq (u(a_i,a_i) - d(2)) + (u(a_j,a_j) - d(2))\,.$$

However, this inequality is inconsistent with $\min\{u(a_i,a_i), u(a_j,a_j)\} > c \geq d(2)$, and therefore even in the case of $c \geq d(2)$, a network with two or more x-components is not an equilibrium network.

Finally, consider the case where there exists an isolated x-player. Assume that player j, $j \neq i$, is in some component C. Similarly to the cases of $c < d(2)$ and $c \geq d(2)$, the condition that player i is not willing to form any link is expressed as follows.

$$0 \geq |C|(u(a_j,a_j)-c)\,, \quad \text{if } c < d(2)\,, \quad (28)$$
$$0 \geq |C| u(a_j,a_j) - (D_j + d(2))\,, \quad \text{if } c \geq d(2)\,. \quad (29)$$

If the component C is an equilibrium component, the utility π_j of player j is nonnegative, i.e.,

$$\pi_j = (|C|-1)(u(a_j,a_j)-c) \geq 0\,, \quad \text{if } c < d(2)\,, \quad (30)$$
$$\pi_j = (|C|-1)u(a_j,a_j) - D_j \geq 0\,, \quad \text{if } c \geq d(2)\,. \quad (31)$$

Because (28) is inconsistent with (30), and (29) is inconsistent with (31), a network with an x-component and an isolated x-player is not an equilibrium. Moreover, because $u(a_j,a_j) > c$, the empty network is also not an equilibrium. Thus, by these facts, the lemma is proven. ■

Proof of Lemma 4

If player i is directly connected with a leaf player j, the utility of player i increases by severing link ij because $c > u(x,x)$. Therefore, equilibrium x-components are leafless, and it must include one or more loops.

When player i interacts with a player of distance L, the condition that link ij is maintained is that $c+\sum_{k=1}^{L} d(k) \leq Lu(x,x)$ holds. This condition is rewritten as

$$c - u(x,x) \leq (L-1)u(x,x) - \sum_{k=1}^{L} d(k),$$

which is the same as condition (6) given in the lemma. Condition (6) means that if the gain arising from the interaction with distant players through a path is larger than the loss, the link cost minus the payoff of the game, of the interaction with an adjacent player directly connected by a link, the link is maintained. Therefore, maintaining an x-component requires all the players in the component to satisfy condition (6). Because the length of a path in a component is at most $\lfloor n/2 \rfloor$, if condition (6) is satisfied for the length $L \in (1, n/2]$ of a loop in an x-component, the x-component can be equilibrium component.

Next, we show that there can exist multiple x-components in an equilibrium network. For sake of simplicity, consider two ring shaped components with $L+1$ players, and they are denoted by C_1 and C_2. Let i and j be players in C_1 and C_2, respectively. After link ij is formed, the variation of the utility of player i is

$$(2L+1)u(x,x) - c - 2\sum_{k=1}^{L+1} d(k). \tag{32}$$

From condition (6), $(2L+1)u(x,x)-c-2\sum_{k=1}^{L} d(k)$ is positive. However, if $d(L+1)$ is sufficiently large, (32) can be negative and then forming link ij results in decrease of the utility of player i. Then, a link between the two components C_1 and C_2 is not formed.

In the other cases such as the case where player i offers to form two or more links with players in C_2, or the case where player i offers to form two links with players in C_2 and one link with another player in C_1, from a similar discussion, one finds that it is possible that the utility of player i decreases, and therefore, in an equilibrium network, there can exist multiple leafless components.

Finally, we consider the condition with respect to the link cost c such that there exists an integer l satisfying condition (6). Because $d(\cdot)$ is a strictly monotone increasing function, if the gain arising from the interaction with distant players through a path is positive, that is, the right hand side of (6) is positive, it is necessary that $u(x,x) > d(2)$. If $c < d(2)$, then from $u(x,x) < c$, one finds $u(x,x) < d(2)$, and there dose not exist an integer l satisfying condition (6). Therefore, $d(2) \leq c$ must be satisfied. When an integer l satisfies condition (6), there exists a loop with $2l+1$ players choosing action x. In the loop, the cost of any player i is $2(c+\sum_{k=1}^{l} d(k))$, and when player i severs a link, the cost changes to $c+\sum_{k=1}^{2l} d(k)$. Because severing a link increases the cost of a player, the following inequality must be satisfied:

$$c \leq \sum_{k=l+1}^{2l} d(k) - \sum_{k=1}^{l} d(k) = RC(2l). \tag{33}$$

Namely, to exist an x-component in an equilibrium network, $d(2) \leq c \leq RC(2l)$ must be satisfied. ■

Proof of Theorem 1

Proof of (1): From Lemma 1, all the players in an equilibrium component choose the same action: α or β. From $c < b$ and Lemma 3, there exists only one α-component or one β-component. For both of the α-component and the β-component, by Lemma 2, the structure of the equilibrium network is determined as described in the theorem.

Proof of (2): Consider the empty β-network. Because there exists no link between players, switching an action from β to α does not change the payoff of a player. Moreover, because forming a link results in decrease of the payoff of a player by $b-c$, such a link is not formed. Thus, any player does not have any incentive to change his strategy, and then the β-network can be an equilibrium network.

As for nonempty networks, from Lemma 1, all the players in an equilibrium component choose the same action: α or β. From Lemma 2 and Lemma 3, by setting $x=\alpha$, one draws the conclusions of (a), (b), and (c) for α-networks, and from Lemma 4, by setting $x=\beta$, one draws the conclusion of (b) for β-networks. In the following, we will prove that a network with both of an α-component and a β-component can be an equilibrium in (b).

Because $d(2) \leq c \leq RC(n-1)$ in (b), from Lemma 4, a β-component can be an equilibrium component. Consider the condition that in an equilibrium network, an α-component with m (≥ 2) players coexists with a ring shaped β-component with $2L+1$ players. If no link between the α-component and the β-component is formed or an α-player and a β-player do not move to the β-component and the α-component, respectively, the network with both of the α-component and the β-component can be an equilibrium.

When a link between an α-player and a β-player is formed, the upper limit of the utility variation of the β-player is

$$mu(\beta,\alpha) - c - (m-1)d(2), \tag{34}$$

and it occurs when the α-component is star shaped. Because (34) is negative if $u(\beta,\alpha) = e < d(2)$, then such a link is not formed. When a β-player changes her action to α, the upper limit of the utility of the β-player is

$$2Lu(\alpha,\beta) - 2c - 2\sum_{k=1}^{L} d(k) + mu(\alpha,\alpha) - c - (m-1)d(2). \tag{35}$$

From $u(\alpha,\beta) = f < e$, the sum of the first, the second, and the third terms of (35) is negative if $e < d(2)$, and then if L is sufficiently large compared with m, (35) can be negative. Therefore, under the above situation, no link between an α-component and a β-component is formed.

Next, consider the case where an α-player moves into the β-component after the α-player changes his action to β, or a β-player moves into the α-component after the β-player changes her action to α. In the former case, it results in decrease of the utility of a β-player originally included in the β-component. Conversely in the latter case, the utility of an α-player originally included in the α-component increases. After the β-player changes her action from β to α, the utility variation of the β-player is

$$2\left(Lb-\left(c+\sum_{k=1}^{L}d(k)\right)\right)-(ma-c-(m-1)d(2)), \quad (36)$$

and if (36) is nonnegative, the β-player does not have any incentive to migrate to the α-component. In order for (36) to be nonnegative, $a \leq Lb-(c+\sum_{k=3}^{L}d(k))$ must be satisfied, and it holds if L is sufficiently large and $b > d(L)$. The condition of ring shaped components in Lemma 2 is compatible with $b > d(L)$, and therefore if L is sufficiently large, (36) can be nonnegative.

From the facts shown above, the network with both of the α-component and the β-component can be an equilibrium.

Proof of (3): From $a < c$, forming a new link results in decrease of the utility of a player, i.e., $u(a_i,a_j)-c<0$ and $u(a_j,a_i)-c<0$, and therefore an empty network is an equilibrium.

As for nonempty networks, from Lemma 1, all the players in an equilibrium component choose the same action: α or β. For $x=\alpha$, from Lemma 4, if there exists an integer $l=l_a$ satisfying condition (6), there exists an equilibrium network with an α-component. Similarly, for $x=\beta$, if there exists an integer $l=l_b$ satisfying condition (6), there exists an equilibrium network with a β-component.

When $c<d(2)$, the right hand side of condition (6) is negative, and then an equilibrium network does not include any component. Namely, an equilibrium network is an empty network, because there does not exist an integer l satisfying condition (6). If there exists an integer l satisfying condition (6), $d(2)\leq c\leq RC(2l)$ must hold. However, because of $2l+1\leq n$, $d(2)\leq c\leq RC(2l)$ implies $d(2)\leq c\leq RC(n-1)$, and therefore when $d(2)\leq c$ and $c>RC(n-1)$, there does not exist such an integer l. Then, an equilibrium network does not include any component. On the contrary, when $d(2)\leq c\leq RC(n-1)$, because this condition is compatible with $d(2)\leq c\leq RC(2l)$, it is possible that there exists a nonempty equilibrium network.

For the coexistence of an α-component and a β-component in the case of $d(2)\leq c\leq RC(2l)$, in a way similar to that of (2), it is shown that a link between the α-component and the β-component is not formed if $u(\beta,\alpha)<d(2)$. Because of $b<a<c$, migration from one component to the other one results in decrease of the utility of a player as shown in the proof of (2). Thus, the network with the α-component and the β-component can be an equilibrium. ■

References

[1] M. O. Jackson and A. Watts, "On the formation of interaction networks in social coordination games", *Games Econom. Behav.*, vol. 41, pp. 265–291, 2002.

[2] S. Goyal and F. Vega-Redondo, "Network formation and social coordination", *Games Econom. Behav.*, vol. 50, pp. 178–207, 2005.

[3] D. A. Hojman and A. Szeidl, "Endogenous networks, social games, and evolution", *Games Econom. Behav.*, vol. 55, pp. 112–130, 2006.

[4] M. Kandori, G. J. Mailath, and R. Rob, "Learning, mutation, and long run equilibria in games", *Econometrica*, vol. 61, pp. 29–56, 1993.

[5] H. P. Young, "The evolution of conventions", *Econometrica*, vol. 61, pp. 57–84, 1993.

[6] G. Ellison, "Learning, local interaction, and coordination", *Econometrica*, vol. 61, pp. 1047–1071, 1993.

[7] E. Droste, R. P. Gilles, and C. Johnson, "Evolution of conventions in endogenous social networks", 2004 [Mimeo].

[8] G. Fagiolo, "Endogenous neighborhood formation in a local coordination model with negative network externalities", *J. Econom. Dynam. Contr.*, vol. 29, pp. 297–319, 2005.

[9] J. Oechssler, "Competition among conventions", *Computat. Mathem. Organiz. Theor.*, vol. 5, pp. 31–44, 1999.

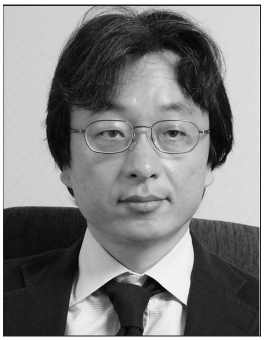
[10] J. C. Ely, "Local conventions", *Adv. Theor. Econom.*, vol. 2, pp. 1–30, 2002.

[11] V. Bhaskar and F. Vega-Redondo, "Migration and the evolution of conventions", *J. Econom. Behav. Organiz.*, vol. 55, pp. 397–418, 2004.

[12] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks", *J. Econom. Theor.*, vol. 71, pp. 44–74, 1996.

[13] V. Bala and S. Goyal, "A noncooperative model of network formation", *Econometrica*, vol. 68, pp. 1181–1229, 2000.

[14] C. Kim and K. Wong, "Network formation and stable equilibrium", *J. Econom. Theor.*, vol. 133, pp. 536–549, 2007.



Ichiro Nishizaki is a Professor at Department of Electrical, Systems and Mathematical Engineering, Faculty of Engineering of Hiroshima University, Japan. His research and teaching activities are in the area of systems engineering, especially game theory, multiobjective decision making and agent-based simulation analysis.

e-mail: nisizaki@hiroshima-u.ac.jp
Department of Electrical, Systems
and Mathematical Engineering
Faculty of Engineering
Hiroshima University
1-4-1, Kagamayama
Higashi-Hiroshima, 739-8527, Japan

Noriyuki Hara is a staff at Technical Center of Hiroshima University, Japan. He works in Division of Radiation Information Registry of Research Institute for Radiation Biology and Medicine of the Hiroshima University.

e-mail: nhara@hiroshima-u.ac.jp
Research Institute for Radiation Biology
and Medicine
Hiroshima University
1-2-3 Kasumi
Minami-ku, Hiroshima 734-8553, Japan

Tomohiro Hayashida – for biography, see this issue, p. 87.

Information for Authors

Journal of Telecommunications and Information Technology **(JTIT)** is published quarterly. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review**. Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

Manuscript. TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The author's JTIT LATEX style file is available:
http://www.itl.waw.pl/publ/jtit/doc/jtitaut.zip

Papers published should contain up to 10 printed pages in LATEX author's style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150–200 words long and the relevant keywords. The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.

Keywords should not repeat the title of the manuscript. About four keywords or phrases in alphabetical order should be used, separated by commas.

The original files accompanied with pdf file should be submitted by e-mail: redakcja@itl.waw.pl

Figures, tables and photographs. Original figures should be submitted. Drawings in Corel Draw and PostScript formats are preferred. Figure captions should be placed below the figures and can not be included as a part of the figure. Each figure should be submitted as a separated graphic file, in .cdr, .eps, .ps, .png or .tif format. Tables and figures should be numbered consecutively with Arabic numerals.

Each photograph with minimum 300 dpi resolution should be delivered in electronic formats (TIFF, JPG or PNG) as a separated file.

References. All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:

[1] Y. Namihira, "Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres", *Electron. Lett.*, vol. 30, no. 3, pp. 262–264, 1994.

[2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

[3] S. Demri and E. Orłowska, "Informational representability: Abstract models versus concrete models", in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314.

Biographies and photographs of authors. A brief professional author's biography of up to 200 words and a photo of each author should be included with the manuscript.

Galley proofs. Authors should return proofs as a list of corrections as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within one week after receiving the offprint.

Copyright. Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the author(s) agree to automatically transfer the copyright for their article to the publisher, if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and all translation rights. No part of the present JTIT should not be reproduced in any form nor transmitted or translated into a machine language without prior written consent of the publisher.
For copyright form see:
http://www.itl.waw.pl/publ/jtit/wskazowki.html

A copy of the JTIT is provided to each author of paper published.

(*Contents Continued from Front Cover*)

INSTYTUT ŁĄCZNOŚCI
PAŃSTWOWY INSTYTUT BADAWCZY

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

tel. +48 22 512 81 83
fax: +48 22 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.itl.waw.pl/jtit