Uniwersytet
Ekonomiczny
w Katowicach

Bartosz Perkowski

Uniwersytet Ekonomiczny w Poznaniu
Wydział Informatyki i Gospodarki Elektronicznej
Katedra Informatyki Ekonomicznej
b.perkowski@kie.ue.poznan.pl

Agata Filipowska

Uniwersytet Ekonomiczny w Poznaniu
Wydział Informatyki i Gospodarki Elektronicznej
Katedra Informatyki Ekonomicznej
agata.filipowska@kie.ue.poznan.pl

MODELLING THE STRENGTH OF RELATIONS IN TELECOMMUNICATION SOCIAL NETWORKS

Summary: Theoretical modelling and empirical studies on networks of various entities have been a subject of a large body of recent research. Methods from the network analysis are applied in such domains as epidemiology, city planning, marketing and social networks.

A social network is a structure consisting of individuals and ties between these individuals. The research on this structure applies social network analysis, with a set of methods, to identify local and global patterns, recognise influential users and examine network dynamics.

Based on the telecommunication data from the Call Detail Records (CDRs), a social network involves individuals (represented by phone numbers) connected by ties based on services used between these individuals.

The paper focuses on the problem of modelling relations (ties) between the telecommunication users in a social network and proposes a method for measuring a strength of a social relation based on CDRs.

Keywords: Social Network Analysis, strength of the relation, telecommunication.

Introduction

The social network is defined as a social structure consisting of individuals that are connected with each other through various types of relations, and between which the information is exchanged on the basis of shared norms and values [Krupski, 2003]. Social networks are used to analyse interactions between individuals as well as an impact of relations on individuals' behaviour. A sample social network may consist of friends of a selected individual, his co-workers, relatives or any other people that are not strictly related to the individual, e.g. friends of friends [Degenne, 1999]. The social network can be modelled with the

use of IT tools and may be represented in a form of a graph.The sociology defines an individual as a person, a being of rational nature of his own existence [Elias, 2008]. The individual is characterised by his independence and possibility of being distinguished among other living organisms. In social networks, an individual may represent, depending on a profile created with the use of social website, a person, a group of people with common interests and an organisation.The connections between individuals in social networks describe various types of relations between people. The types of relations distinguished in psychology are: friendship, love, marriage, kinship, co-workers and neighbours [Argyle, 2002]. The connections are crucial in forming a social structure, thus it is important to model the social nature of the relation. The social relation is defined as an interaction between individuals or groups, taking place in accordance with accepted practices and schemas [Znaniecki, 2011]. The individuals' actions can be classified as socially positive, when they interact to achieve a favourable outcome, and socially negative when at least one of these individuals acts contrary to applied practices.The relations between individuals, as they play a key role in the social structure formulation, are the subject of the Social Network Analysis. The type of a relation is represented in a social network with its strength: the stronger the bond between two users, the higher the value that indicates the strength of a relation. For example, spouses are in a stronger relation than friends, who in turn are closer to each other than acquaintances. The aim of the Social Network Analysis is to understand events that occur in social networks, especially including the existence of interdependence between individuals and the analysis of the impact of various factors on the structure of the network.The telecommunication social network emerges from the connections between individuals performed with the use of telecommunication services. As in other social networks, a key role plays the type of relation, that describes a strength and qualifies the connection between individuals. The problem that occurs in telecommunication social networks is related with the identification of attributes of a connection that influences the closeness of a relation. The presentation of approaches to the Social Network Analysis and to the formulation of relations between individuals enables achieving the main goal of the paper. This aim is to identify the connection's attributes and to introduce a method for modelling a strength of a relation in telecommunication social networks.The structure of the paper is as follows. The first section contains the review of the social networks domain. It presents the motivation to perform research in the field of social networks with examples of social networks. The second section concerns telecommunication social networks. The analysis of the data types processed by telecommunication providers is performed and the levels of telecommunication

networks' analysis are specified. Moreover, this section contains the description of characteristics of Call Detail Records and the analysis of methods used for modelling the relations. Based on the CDR analysis, the third section focuses on the data model used in the study and the description of the assumptions made regarding the attributes affecting the social nature of a relation. On this basis, a method for measuring a strength of a relation has been proposed and tested over relations between individuals. Finally, the comparison of the selected methods for measuring a strength of the relation has been performed. The last section summarizes and describes the impact of the presented solution on the social network structure. In addition, it proposes further work over the method.

1. The analysis of social networks and connections between individuals

The history of the Social Network Analysis goes back to 1887, when researchers started to analyse the process of formulation of communities based on diverse relations between individuals. The community can be described as a group of individuals (being a part of a social network), where connections between individuals inside the group are stronger than connections with individuals outside the group. The results of the analysis revealed, that communities are formed based on the personal relations that include emotions and intimacy, and on formal relations arising from the organizational structure [Tönnies, 1887]. Further work concerned the impact of various social and psychological factors on formation of relations. The sample research showed, based on the observations of children in a classroom, that sex is an important factor while dividing network of individuals into smaller communities [Moreno, 1943]. Moreover, the social exchange theory also affects the formation of relations. The theory posits that human relationships are formed by the use of a subjective cost-benefit analysis and the comparison of alternatives [Blau, 1964]. Also formal relations impact the creation of communities. Researchers examined the link between the organizational structure and the closest contacts of individuals [White, 1970; Magee, 2008].

The Social Network Analysis became more popular with the evolution of social media. Social websites are services used to establish contacts between individuals and building relations between them based on common interests and activities. The social network created by users of a social website consists of a representation of each individual called a user profile, its social relations and a variety of additional services.

Facebook[1] is a worldwide social network that has 1.44 billion active users monthly (as of 31 March 2015)[2]. Each user has a possibility to connect to others by making online friends. A Facebook user can also be an organisation, which is represented as a fanpage. In that case, a relation is formed, when any user likes the fanpage. Each user can contact with another by one of two types of communication: private – a very similar to e-mail, and public, where users have a possibility to comment on entries, post photos, video footage and links to websites [Valenzuela, 2009]. Depending on a type of a relation, a Facebook social network reflects the network of friends or a network of users that are in contact.

The Social Network Analysis is the study of human interactions using the network theory. A part of research regarding relations in social networks concerns the "small-world experiment" conducted by S. Milgram in 1967. He examined the average path length in social networks of people in the United States and, based on the results, suggested that a human society is a small-world-type network characterized by short path-lengths [Milgram, 1967]. The experiment is associated with the "six degrees of separation" theory. This theory was verified for Facebook in 2011, showing that the average number of degrees of separation between any pair of users is 3.47.

An example study of the impact of sociological factors on the social network structure is related to the theory of homophily. It describes the tendency of individuals to associate and bond with similar people. Numerous studies have shown that the similarities between individuals increase the likelihood of establishing a connection. These similarities may be associated with age, gender, class and organisational role [McPherson 2001]. Individuals, who are in a relation arising from the similarities, share common characteristics, such as beliefs or values that strengthen the relation. Therefore, the Social Network Analysis focuses not on the individual and its profile, but on the relations that occur between individuals or communities.

The latest research shows that the quality of Social Network Analysis is higher, when the nature of the relation is addressed [Liben-Nowell, 2003]. Many research works regarding social networks describe the relation in a binary form, what causes the nature of the relation to disappear or to be shown in a rudimentary way. Treating all relations as binary (existing/no existing) limits the analysis of existing social networks [Xiang, 2010], and in particular, highlighting pairs of individuals who are in a strong relationship (e.g. family) or couples with

---

[1] https://www.facebook.com/.
[2] http://newsroom.fb.com/company-info/.

a weaker connection (e.g. friends). The implementation of a measure for a strength of a relation that involved the valuation of connections in social networks, allowed better reproduction of relations between individuals and has initiated research on the importance of a strength of relations in the social network's structure. It has been proved that not only strong ties are important, but also weak ones can be crucial, especially in case of relations that connect individuals from different communities [Granovetter, 1973]. The present works on the Social Network Analysis focus on determining a strength of relations by including their social nature. The recent analyses of social media like Facebook and Twitter focus on predicting the tie strength based on predictive variables that could represent the social nature of the relations: predictive intensity variables (e.g. wall words exchanged, inbox messages exchanged, friend's status updates), intimacy variables (e.g. individual's number of friends, wall intimacy words, appearances together in a photo), duration variables (e.g. days since first communication), reciprocal services variables (e.g. links exchanged by wall post, applications in common), structural variables (e.g. number of mutual friends, groups in common), emotional support variables (e.g. wall and inbox positive emotion words) and social distance variables (e.g. age, educational, political difference) [Gilbert, 2009].

2. Social networks in telecommunication

Research results on social networks are transferred to different domains of social life. One of them are social networks resulting from relations between individuals in telecommunication networks. The telecommunication industry creates, stores and processes a huge amount of data. This data includes call logs that describe the usage of services by subscribers, network data related to the status of hardware components and software, and customer data describing profiles of telecommunication subscribers [Weiss, 2005].

Assuming, that the telecommunication social network is a derivative of communication that occurs with the use of telecommunication services, the customers' analysis is carried out on telecommunication data applying techniques taken from the Social Network Analysis. The classic approach to represent the social network based on call logs is to include the number of connections between individuals as well as duration of phone calls, and combine these with users profiles created based on data from other sources, such as Facebook [Puspa, 2012]. However, it should be noted that acquiring this data may be costly as well as combining data from multiple sources can be a very time-consuming task, re-

quiring more computing power than in case of data from a single source. Thus, the expenditure on processing and integration of data may be higher than expected return from the analysis.

Call Detail Records are call logs that store the data produced by a telephone or any other telecommunication equipment. CDRs contain the details of a phone call or other communication transaction (e.g. a short text message), which passes through the device. They describe the metadata of each telecommunication event, but do not contain the content of the event. In a simplified case, for a phone call, the description of an event contains the phone numbers of a caller and receiver, the start date and time of an event and the duration. Current CDRs expand this description by adding for e.g. the identifier of a connection, the identifier of a Base Transceiver Station (BTS) used by the caller, an identifier of a service used, a status of a connection and an error code, if occurred.

The research carried out on the Call Detail Records proved that it is possible to draw inferences about personality of an individual based on his behaviour associated with the mobile phone usage [de Oliveira, 2011; de Montjoye, 2013]. Using CDRs, the characteristics of a user can be compared with his psychological description derived through mirroring the variables from CDRs on qualitative variables. This can help in determining potential differences between personality types defined in a psychological five-factor model [Matthews, 2003]. Analysis of the correlation of defined variables (e.g. basic phone use, spatial behaviour or regularity) with personality types, showed a high convergence. On this basis, it can be assumed that by taking into account relevant attributes of communication between individuals, it is possible to measure a strength of the relation with regard to the social nature of the relation.

The research problem related to the analysis of social networks in telecommunications concerns taking into account only quantitative attributes of communication between users to model these networks. As it was presented on the Facebook example, a strength of ties should not only be a result of combining communication types, number of interactions and their duration, but also should include the social nature of this relation. In order to propose a novel method for measuring a strength of a relation based on the telecommunication data, a few questions have to be answered: What attributes of communication have an impact on a strength of ties between individuals in a social network? How to reflect the nature of relation with attributes of edges in a graph? What is the data model that should be used to represent social relations in a graph database?

3. Proposed method for measuring the strength of the relation

The complexity of the problem of measuring strength of a relation in tele-communication social networks relates to the sociological and psychological approach to interpersonal relations. It is necessary to adapt communication models defined in psychology that describe the direction of communication and its impact on a strength of a relation [Morreale, 2007]. Moreover, the sociology defines dynamic attributes of communication, which include a location, user behaviour patterns, the closeness of contacts and the mutuality [Choi, 2013]. Combining quantitative with dynamic quantitative attributes of a relation, it is possible to describe strength of relations between individuals referring to social aspects of relations. The key element of the method is describing the form of communication, resulting from the service used by an individual to communicate. The sociological literature indicates the superiority of the verbal communication over the non-verbal with the ability to transfer information in a more understandable way, by enriching the information with expressed emotions. On this basis, it can be assumed that phone calls are more important than text messages. Another attribute of the social relation is the duration of a phone call, indicating an increasing closeness of a relation between individuals due to increasing call duration. Moreover, individuals that form a strong bond, tend to contact one another, thus the frequency of contacts has an impact on strength of a relation. Missed calls may indicate the unavailability of the recipient, but also the recipient's reluctance to continue a relation with a caller. Therefore, it can be assumed that the frequency of answered calls strengthens the link, and the rate of missed calls has a negative impact on the strength. The final assumption related to the social nature of a relation is the time of day, during which the connections occur. The day can be split into two parts: an office time and a private time. As a result, a specific nature of the relation can be selected depending on the time in which the communication takes place.

To apply the method of measuring strength of a relation in a graph structure that represents the telecommunication social network, it is necessary to design a data model consistent with the specified assumptions. The data model is presented in . The node represents a phone number from Call Detail Record. Each node has been assigned with a unique identifier that allows for a quick phone number search in the database, the identifier of the owner and the number of the Base Transceiver Stations used. The edge represents the type of a connection between a pair of nodes. Types of relations have been defined according to the service used for communication, time of day and the status of connection (answered

or missed). Based on the combination of the connection's attributes, six types of relations were listed: answered calls in the office time (CMA), missed calls in the office time (CMNA), answered calls in the private time (CEA), missed calls in the private time (CENA), text messages sent during the office time (SM), and text messages sent in the private time (SE).
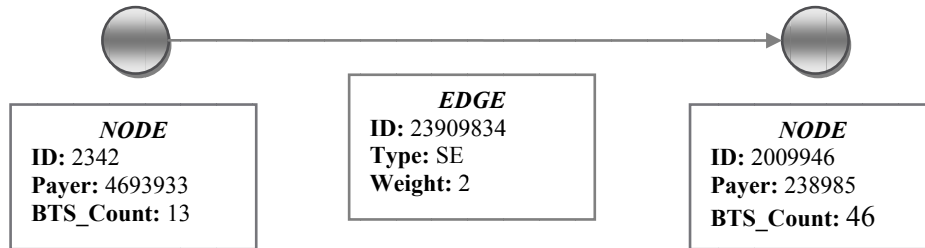

Fig. 1. Data model with attributes

The initial test of the proposed measure was performed over a randomly chosen set of numbers from a three month Call Detail Records that contain 33.3 million nodes and 93.8 million edges. For all contacts of a selected user (node), the experiment concerned calculation of a basic measure (baseline, quantitative) and the new measure (with emphasis on social aspects). In order to compare the selected measures, all results had to be normalized. The comparison of normalized measures for a selected user is presented in Fig. 2.
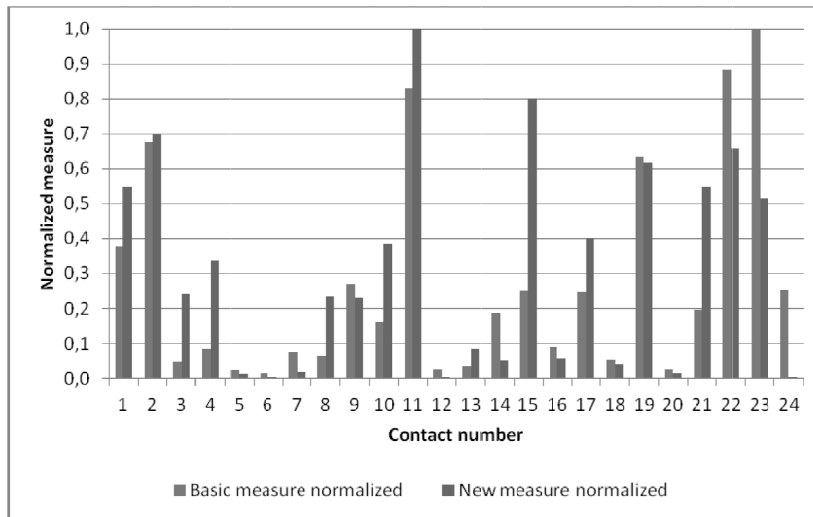

Fig. 2. Comparison of normalized measured for strength of a given relation of a selected user

The initial version of the method gives a first overview on shaping strength of the relation with the use of qualitative attributes of the connections. According to the assumptions, the primary factor in forming a relation is a phone call. The mutual relations, for which there are a low percentage of missed calls, are favoured. The duration of calls results in a logarithmic increase of strength of the relation, in contrast to the basic measure, which increases strength in a proportionate manner. Moreover, the function of text messages is to strengthen the relation if it occurs. These primary results show that qualitative attributes have a strong impact on the strength of a relation. By strengthening the mutual relations between individuals and weakening the one-sided relations, the measure provides a better representation of relations, in which individuals tend to contact with others. The detailed statistics for selected contacts of this user are shown in Table 1.

Table 1. Basic statistics of a phone usage for selected contacts of a user

Number	Outgoing calls	Incoming calls	Outgoing answered calls	Incoming answered calls	Text messages sent	Text messages received	Sum of duration of outgoing calls	Sum of duration of incoming calls
2	20	29	19	26	0	0	3507	4676
4	19	19	15	14	0	0	1023	983
11	52	84	33	63	40	48	3564	6106
15	33	23	22	18	2	7	2203	2468
22	56	49	55	30	0	0	8413	4010
24	8	1	8	0	0	0	2594	0

Conclusions

The goal of the paper was to outline the problem of modelling strength of relations in social networks, with a particular emphasis on telecommunication social networks. The analysis of the literature in the domain of social networks and interpersonal communication indicated attempts to solve the problem of measuring a strength of relations between individuals in networks like Facebook, by introducing qualitative characteristics resulting from an individuals' behaviour and a submitted content. It has also been shown that the problem of measuring a strength of relations applies to telecommunication networks, for which the customer analysis is crucial in the context of marketing activities, fraud detection and customer migration.

The current state of research made it possible to propose the assumptions for the method that includes in relation strength the social nature of a relation resulting from the connections created with the use of telecommunication services. These assumptions enable the identification of the social nature of relation described in the literature, based on the information stored in Call Detail Records. It is important to describe the behaviour of individuals not only on the basis of quantitative attributes, such as the number of calls and their total duration, but by enriching these attributes with behavioural patterns resulting from the frequency of connections or the mutuality of relation.

The second result of presented work is the data model for the graph database, based on the information stored in Call Detail Records, and the identification of a problem related to processing of large data sets. The proposed data model has two main functions: it provides the ability to implement the algorithm for assessing strength of a relation with the designed method, and supports the efficient data load and processing. Having the data model implemented, it was possible to import the data from CDRs into a graph database and run initial tests for the new method of measuring the relation's strength. The results show significant differences between tested methods.

As a part of further work, an empirical research to verify the proposed assumptions for the method will be performed with the use of the online questionnaire. The respondents will evaluate the importance of each attribute of a relation and its impact on strength of a relation. The results of the survey will be used to verify statistical hypotheses concerning the relevance of each attribute. Based on the results of verification, the method will be modified and tested over a random sample of users and relations, as it was presented in Section 3. The calculated measures will be used to verify a hypothesis on a similarity of the relation strength for both methods (basic and with social flavour). In addition, case studies will be conducted to describe a selected relation and relation strength will be measured for selected pairs of individuals.

References

Argyle M. (2002), *Psychologia stosunków międzyludzkich*, Wydawnictwo Naukowe PWN, Warszawa.

Blau P.M. (1964), *Exchange and Power in Social Life.* Transaction Publishers, Livingston Campus, NJ.

Choi J., Heo S., Han J., Lee G., Song J. (2013), *Mining Social Relationship Types in an Organization Using Communication Patterns*, Proceedings of the 2013 conference on Computer supported cooperative work, ACM, New York, NY.

Degenne A., Forse M. (1999), *Introducting Social Networks*, Sage Publications, London.

Elias N. (2008), *Społeczeństwo jednostek*, Wydawnictwo Naukowe PWN, Warszawa.

Gilbert E., Karahalios K. (2009), *Predicting Tie Strength With Social Media*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 211-220.

Granovetter M.S. (1973), *The Strength of Weak Ties*, "American Journal of Sociology" Vol. 78, No. 6, pp. 1360-1380.

Krupski R. (2003), *Zarządzanie strategiczne. Koncepcje. Metody*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Liben-Nowell D., Kleinberg J. (2003), *The Link Prediction Problem for Social Networks*, International Conference on Information and Knowledge Management.

Magee J.C., Galinsky A.D. (2008), *The Self-Reinforcing Nature of Social Hierarchy: Origins and Consequences of Power and Status*, IACM 21st Annual Conference Paper.

Matthews G., Deary I.J., Whiteman M.C. (2001), *Personality Traits*. Cambridge University Press, Cambridge.

McPherson M., Smith-Lovin L., Cook J.M. (2001), *Birds of a Feather: Homophily in Social Networks*, "Annual Review of Sociology", Vol. 27.

Milgram S. (1967), *The Small-World Problem*, "Psychology Today", Vol. 2, pp. 61-67.

Montjoye de Y.-A., Quoidbach J., Robic F., Pentland A. (2013), *Predicting Personality Using Novel Mobile Phone-Based Metrics*, Social Computing, Behavioral-Cultural Modeling and Prediction.

Moreno J.L., Jennings H.H., Stockton R. (1943), *Sociometry in the Classroom*, "Sociometry" Vol. 6, No. 4, pp. 425-428.

Morreale S.P., Spitzberg B.H., Barge J.K. (2007), *Komunikacja między ludźmi. Motywacja, wiedza i umiejętności*, Wydawnictwo Naukowe PWN, Warszawa.

Oliveira de R., Karatzoglou A., Concejero Cerezo P., Lopez de Vicuña A.A., Oliver N. (2011), *Towards a Psychographic User Model From Mobile Phone Usage*, CHI '11 Extended Abstracts on Human Factors in Computing Systems.

Puspa, Shobha G. (2012), *Social Network Analysis for Churn Prediction in Telecom Data*, "International Journal of Computer & Communication Technology", Vol. 3, No. 6, pp. 128-135.

Tönnies F. (1887), *Gemeinschaft und Gesellschaft*, Fues's Verlag, Leipzig.

Valenzuela S., Park N., Kee K.F. (2009), *Is There Social Capital in a Social Network Site? Facebook Use and College Students' Life Satisfaction, Trust, and Participation*, "Journal of Computer-Mediated Communication", No. 14(4), s. 875-901.

Weiss G.M. (2005), *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Data Mining in Telecommunications, Kluwer Academic Publishers, pp. 1189-1201.

White H.C. (1970), *Chains of Opportunity: System Models of Mobility in Organizations*. Harvard University Press, Cambridge, MA.

Xiang R., Neville J., Rogati M. (2010), *Modeling Relationship Strength in Online Social Networks*, Proceedings of the 19th International Conference on World Wide Web.

Znaniecki F. (2011), *Relacje społeczne i role społeczne: niedokończona socjologia systematyczna*, Wydawnictwo Naukowe PWN, Warszawa.

MODELOWANIE SIŁY POWIĄZAŃ W TELEKOMUNIKACYJNYCH SIECIACH SPOŁECZNOŚCIOWYCH

Streszczenie: Modele teoretyczne oraz badania empiryczne skoncentrowane wokół sieci tworzonych dla różnego rodzaju jednostek są w ostatnich latach przedmiotem wielu badań. Metody, pochodzące nie tylko z matematyki, ale także z psychologii i socjologii, mają zastosowanie w takich dziedzinach, jak epidemiologia, urbanistyka, marketing czy sieci społeczne. Sieć społeczna jest strukturą społeczną składającą się z jednostek, które są połączone ze sobą poprzez różne rodzaje relacji (od przypadkowych spotkań po więzi rodzinne) i pomiędzy którymi dochodzi do wymiany informacji na zasadach wspólnie wyznawanych norm i wartości. W socjologii sieć społeczną definiuje się jako grupy ludzi połączonych ze sobą. Badania tej struktury związane są z analizą sieci społecznych, wykorzystującą zestaw metod umożliwiających identyfikację lokalnych i globalnych schematów, rozpoznawanie wpływowych jednostek w sieci oraz badanie dynamiki sieci. Na podstawie danych telekomunikacyjnych pochodzących z dzienników połączeń (CDR) może zostać utworzona telekomunikacyjna sieć społecznościowa, w której jednostki reprezentowane za pomocą numerów telefonów są ze sobą powiązane poprzez zdarzenia zachodzące między nimi, z wykorzystaniem różnego rodzaju usług telekomunikacyjnych.

W artykule przedstawiony został problem modelowania powiązań pomiędzy jednostkami należącymi do telekomunikacyjnej sieci społecznościowej oraz zaproponowana została metoda mierzenia siły relacji społecznych między tymi jednostkami na podstawie dzienników połączeń.

Słowa kluczowe: analiza sieci społecznościowych, siła relacji, telekomunikacja.