

# Mathematical Model of Analysis and Optimal Harmonisation of Data for Preparation of Broadband Infrastructure Development Project in the Republic of Croatia: Contribution to Achieving Objectives of the Digital Agenda for Europe

Matej Požarnik<sup>1</sup>, Tea Taras<sup>1</sup>, Aleš Kranjec<sup>1</sup>, Lea Robič Mohar<sup>1</sup>

<sup>1</sup>ProFUTURUS d.o.o., Maribor, Slovenia Correspondence: Lea Robič Mohar, ProFUTURUS d.o.o., 2000 Maribor, Slovenia.

Received: August 1, 2016Accepted: August 10, 2016Online Published: August 19, 2016doi:10.11114/bms.v2i3.1823URL: http://dx.doi.org/10.11114/bms.v2i3.1823

## Abstract

Next Generation Access projects of developing broadband infrastructure, co-financed by the EU, require preparation and implementation of complex preparatory actions that are crucial for their successful implementation. Their implementation within member states contributes directly to achieving goals, defined in the Digital agenda for Europe. Preparation work on projects requires complex analyses and processing of governmental data on potential users of the new network. A mathematical model was developed within this study. Also, the Optimal Data Harmonization algorithm was developed, which incorporates the pointed mathematical model. Large amounts of projects' input data were also processed with the Optimal Data Harmonization algorithm. Research showed that data processing without Optimal Data Harmonization algorithm is the key reason for inadequate preparation of the project because of imprecise definition of numbers and geolocations of potential users. Without applying such model, projects don't achieve the target coverage of broadband infrastructure of the next generation access and therefore don't contribute to the Digital agenda for Europe goals. Processing data with Optimal Data Harmonization algorithm, which is a beyond the state-of-the-art model, ensures a high-level harmonisation of national data. With these results, one can provide optimal coverage of eligible potential beneficiaries in the project. Projects in which data were processed with the Optimal Data Harmonization algorithm contribute to reaching the Digital agenda for Europe goals. The study also recommends the establishment of national central database of geolocations of the potential beneficiaries, as it would standardize the input data in all future projects.

Keywords: broadband, data management, optimal data harmonization algorithm, cohesion policy, digital agenda for Europe

## 1. Introduction

The European Union (EU) undertakes substantial efforts to equalize the development of its regions. EU cohesion policy is defined by the strategic document Strategy Europe 2020 (European Commission, 2010b), which is implemented through objectives; one of the key Europe 2020 flagships, used for this study being the Digital Agenda for Europe (DAE).

DAE defines the following objectives (European Commission, 2010a):

- Achievement of full broadband coverage of the minimum speed  $\geq$  30 Mbit/s,
- Usage of broadband minimum speed of  $\geq$  100 Mbit/s in at least 50% of households.

Implementation of projects of developing Next Generation Access (NGA) infrastructure in less developed areas requires high investments, which are usually not profitable for the operators. Considering the above, the EU co-finances such projects from structural funds, while the member states adopt national programs for the development of the NGA, which are in line with the EU guidelines, and at the same time also define national guidelines of implementation of such projects (European Commission, 2013).

In the Republic of Croatia (RH), a national indicative program (ONP) for the development of broadband infrastructure was adopted for areas where there is insufficient commercial interest for investments (Government of the Republic of

Croatia, 2016). ONP regulations define areas that are eligible for co-financing, which is determined by the process of determining the quality of internet access to the individual geolocation, or by "mapping". Geolocations without the NGA are referred to as "white areas" and are as such the only ones eligible for co-financing by the EU funds (Government of the Republic of Croatia, 2016).

## 1.1 Definition of the Problem

Working on various projects in RH, some inconsistency and incompatibility of information on locations of potential users was established, which however serves as the basis for the "mapping" procedure. Furthermore, a mismatch in their relation to data on geolocations of house numbers exists. Also, it was found that the state-of-the-art data processing method results in anomalies in the sense of imprecise definition of geolocations of all eligible potential beneficiaries, or geolocations of new connections. Such inadequate base consequently causes lower than planned NGA coverage, which also deviates the real financial values of the project.

Given that the number of potential users in "white areas" serves as the base for defining the number of connections, eligible for EU co-financing, inaccurate determination of geolocation and the number of connections causes unjustifiable exclusion of a certain number of eligible potential beneficiaries from the project. Because of their omission, it will be impossible to achieve the DAE objectives and successfully implement the EU cohesion policy.

Within this study, a mathematical model was developed that is incorporated in the Optimal Data Harmonization (ODH) algorithm (beyond state-of-the-art). It provides an adequate analysis and processing of government data, as well as precise definition of all eligible beneficiaries as the basis for the "mapping" process.

During research, the ODH algorithm was the integral part of the broadband deployment projects. By implementing it into the preparatory activities of the project, we obtained input data which defined the numbers and geolocations of potential users with high precision. This provides a high level of coverage of NGA and thus directly contributes to achieving the DAE objectives.

#### 1.2 Hypothesis

Problems, defined in the introductory part, initiate the following question: whether and how does the mismatch of state records affect the implementation of the EU cohesion policy in the EU member states?

This question is the basis for setting the null hypothesis (Ha) and the alternative hypothesis (Hb):

Ha: the mutual incompatibility of data and records, necessary for preparation of projects for the development of broadband infrastructure, co-financed by the EU, prevents the achievement of goals, as set by the DAE, and hence successful implementation of the European cohesion policy.

Hb: the mutual incompatibility of data and records, necessary for preparation of projects for the development of broadband infrastructure, co-financed by the EU, has no influence to the achievement of goals, as set by the DAE, and hence successful implementation of the European cohesion policy.

#### 2. Literature survey

## 2.1 European Cohesion Policy and NGA

Successful internal cohesion is a requirement for effective achievements of the EU objectives in terms of foreign policy (da Conceição-Heldt and Meunier, 2014). According to Bachtler and Turok (2013), crucial to the implementation of cohesion policy are the Cohesion and Structural funds. Reggi and Scicchitano (2014) conclude that a possible reason for the lack of cohesion is the lack of a proper analysis of the regional context and conditionally, the policy decisions (Reggi and Scicchitano, 2014). Research results of McCann and Ortega-Argilés (2015) indicate the necessity of relying the European cohesion policy on the logic of smart specialization, which represents the foundation of information and communication technology (ICT).

The EU records a significantly higher level of development in regions with capital cities, or in urban areas of the EU (Cuaresma, Doppelhofer, and Feldkircher, 2014). Auhors emphasize the importance of education and ICT, in particular in the field of broadband coverage of the new generation access (NGA). Marcut (2014) analyses social achievements of the EU, putting them in relation with the network society. Broadband infrastructure is the basic prerequisite for the development of the "Smart city" concept, the positive aspects of which are demonstrated by Carter (2013).

Daly (2012), on the other hand, concludes that the strategy Europe 2020 lacks a coherent model of social development, while McCormick (2013) finds that the EU has made too little progress in the creation of a single digital market because of a limited availability of NGA which is slowing down the economic growth.

Van der Wee, Verbrugge, Sadowski, Driesse, and Pickavet (2015) state that the availability of the NGA prevents depopulation of less developed areas. In addition, the authors demonstrate indirect positive effects of NGA, related to

e-Business and e-Government, and consequently additional incentives for investments into NGA networks (Van der Wee et al., 2015). Furthermore, broadband NGA infrastructure reduces costs for health care (Currie, Philip, and Roberts, 2015) and promotes life-long learning and distance learning (Rampersad and Troshani, 2013). The Ministry of Maritime Affairs, Transport and Infrastructure of the Republic of Croatia (2014) has acknowledged the positive effects of NGA through a 4.5% average increase of the highest level of education as the consequence of investing into NGA infrastructure and using its possibilities.

Long-term benefits were examined by Briglauer and Gugler (2013) who estimate that intensive increase of using NGA to internet results in increased GDP (between 0.47% and 1.38%), while according to the study of the World bank, a 10% increase of investments into NGA infrastructure results in an 1.21% increase of the GDP in less developed countries (Qiang, Carlo, and Kaoru, 2009).

## 2.2 State Interventions and Data

As concluded by Cambini, Briglauer, and Grajek (2015), public support for development of NGA infrastructure is an excellent way for encouraging investments. Gulati and Yates (2012) record a higher level of prevalence of NGA infrastructure in countries with a higher level of financial investments into ICT technologies, with simultaneous effective national regulatory function. On the other hand, in developing countries, a competitive telecommunication sector is the leader of the expansion of broadband access, while the national regulation would have a negative effect (Gulati and Yates, 2012).

Groot (1997) concludes that a sustainable harmonization of geospatial data depends on the political responsibility and integrity, and a local regulatory environment which must be aligned with the national and supranational level. Villa, Molina, and Gomarasca (2013) analysed scenarios of geodata harmonisation, recording positive results according to HUMBOLT access.

## 3. Methods

#### 3.1 Sample

This research included projects for development of broadband infrastructure that cover 370 872 potential beneficiaries in the Republic of Croatia (RH). The number of potential beneficiaries that are included in this research is equivalent to an average of 21816 per project. These projects were implemented in areas of the groups I - V according to the development index, as suggested by the Ministry of regional development and EU funds. This survey used data on private households, companies, artisans/craftsmen, public users and data on geolocation of the individual house number.

#### 3.2 Instruments and Procedure

Working on various projects in Croatia, individual phases and results of data processing were studied as the foundation for defining physical and financial basis for implementing projects.

In RH, there is no base (directory) of geolocations of potential beneficiaries; there is only a geolocation of house numbers. Therefore, for the purpose of determining the geolocation of potential beneficiaries, access to data on their location was obtained. The latter is integrated with the unique data on geolocation of all house numbers in the project area.

Data was obtained from the following relevant state resources:

- State Geodetic Administration (DGU).
- Federal Bureau of Statistics.
- Ministry of the Interior of RH.
- Financial agency.
- Ministry of Business and Trade RH.
- Project owners (counties and cities/municipalities).

Obtained data was processed, organised and cleaned. All detected anomalies were exploratory analysed and eliminated. Locations of potential beneficiaries were paired with geolocations of DGU, applying state-of-the-art method and ODH algorithm (beyond state-of-the-art). The value of investment was calculated, based on data, processed with state-of-the-art method and those data, processed with ODH algorithm.

#### 3.2.1 State-of-the-art

Processing data through state-of-the-art method was done using obtained data in its original form, its analysis, organisation and cleaning. Taking into consideration the fact that data is provided from centralized registers of the government, it was presumed they were originally compatible.

3.2.2 Beyond State-of-the-art -ODH Algorithm

With help of incorporated mathematical model, the ODH algorithm enables optimal mutual harmonisation of government data with the goal to precisely assess the number of potential beneficiaries of the new NGA network.

Schematic representation of data processing and results, obtained from state-of-the-art method and through help of ODH algorithm is shown parallel in Figure 1.



Figure 1. ODH algorithm

Note: B = the number of eligible potential beneficiaries;  $B_0 =$  initial number of eligible potential beneficiaries; n = difference to the actual number of eligible potential beneficiaries;  $B_1 =$  the number of eligible potential beneficiaries after the ODH algorithm.

ODH algorithm is biphasic:

1. Equation of data on location of potential beneficiaries with unique data on locations of house numbers.

Equation is done through a mathematic procedure that is shown with the following formula:

$$L_{PB}(c,t,s,h) = L_{HN}(c,t,s,h) \tag{1}$$

Whereby the following applies:

 $L_{PB} = location of potential beneficiary,$ 

 $L_{HN} = location of the house number;$ 

c = county,

t = town,

s = street,

h = house number.

2. Defining the number of potential beneficiaries on individual location.

Defining the above is done with help of a mathematic procedure that can be shown with the following formula:

$$GL_{PB} = x, y, z_{HTRS96}(L_{PB}) \tag{2}$$

$$N_{PB}(GL_{PB}) = \sum N_{PB}(L_{PB}), \ L_{PB} = GL_{PB}$$
(3)

Whereby the following applies:

 $GL_{PB}$  = geolocation of potential beneficiary,

 $L_{PB}$  = location of potential beneficiary,

 $x, y, z_{HTRS96}$  = actual physical location in a coordinate system HTRS96,

 $N_{PB}$  = number of potential beneficiaries.

#### 4. Results

With the completion of the first phase of ODH algorithm, approximately 42.13% of changes of original data were noticed within research.

With the completion of the second phase of the algorithm, approximately 90.56% of paired data on locations of potential beneficiaries with geolocations of house numbers was registered. The least paired data are observed in data on public users (79.47%); the highest rate being observed in data that define the geolocation of private households (92.01%).

Figure 2 shows the average level of pairing relevant data in all projects that participated in research.



Figure 2. Levels of pairing data on geolocations

Note: These data are aggregated from all researched projects, and shown in average values.

Figure 2 shows a total of 45.87% larger pairing of locations of potential beneficiaries with the geolocation of house numbers.

Table 1 shows natural project inputs (the number of paired locations of potential beneficiaries), processed with state-of-the-art method and those processed through ODH algorithm.

Table 1. Natural inputs, processed with state-of-the-art method and ODH algorithm

	State-of-the-art	ODH algorithm
Private households	11871	16550
Companies	852	1771
Artisans/craftsmen	765	1307
Public users	56	129
Total	13544	19757

Note: The data are aggregated from all researched projects and shown in average values.

Table 1 shows that the ODH algorithm results in defining 39.42% more households, 107.86% more companies, 70.85% more artisans & craftsmen and 130.36% more public users as compared to the state-of-the-art method of data processing.

Based on in-kind data, collected and processed as part of the research, financial aspects of the project were defined as shown in Table 2.

Table 2. Value of investment based on the type of technology in EUR

	State-of-the-art	ODH algorithm
VDSL (FTTC)	6 645 589	9 694 101
GPON (FTTH P2MP)	17 282 144	25 209 932
FTTH P2P	19 954 827	29 108 647
Cable access (DOCSIS, HFC)	8 072 224	11 775 172
LTE (4G)	15 819 392	23 076 176
FTTC / FTTH	15 566 571	22 707 379

Note: Data are aggregated from all research projects and presented in average values.

The value of investment is calculated based on the estimated value of the connection; relative deviations therefore depend on the variable price of the connection, depending on technology.

#### 5. Discussion and Conclusion

Research results indicate a high level of non-compliance of data from public records in Croatia, processed through state-of-the-art method. Analysis of financial data shows a strong deviation of the investment value due to the

sensitivity of the topic - project of development of NGA broadband infrastructure.

The mutual incompatibility of data and records and their state-of-the-art processing in projects' preparation resulted with the average of 13544 potential beneficiaries per project. As the research included an average of 21816 potential beneficiaries per project, the results show that 8272 of them or 37.92% would not be covered. Considering at least one DAE target (achievement of full broadband coverage of the minimum speed  $\geq$  30 Mbit/s), the lack of covering 37.92% of beneficiaries will fail on achieving the full broadband coverage.

The above discussed results confirm the defined Ha and reject the Hb: omitting a certain number of eligible users does not achieve the targeted coverage with NGA broadband access, as defined within DAE. Since reaching this objective is one of the tools of the cohesion policy, this state of affairs hinders its successful implementation.

The unique contribution of this research is to develop an ODH algorithm that includes a mathematical model of precise definition of numbers and geolocations of potential beneficiaries in broadband deployment projects. Implementing the said algorithm ensures usage of coordinated information. Consequently, a large number of eligible potential beneficiaries are included in projects.

ODH algorithm defined 90.56% of accurate geolocations of all potential beneficiaries. Defining the geolocation and the number of potential beneficiaries with 100% accuracy means pairing all locations with a unique geolocation of house numbers. Results show that the ODH algorithm achieved a total of 45.87% higher pairing of data than compared to the state-of-the-art method. Each increase in the coverage of NGA due to its ambiguous benefits exponentially increases the level of economic and social development of the area. Therefore, the realized increase of percentage of data pairing has a considerable impact on success of the project from the point of its impact on implementation of cohesion policy.

The problem of inadequate definition of the geolocation and the number of potential beneficiaries is related to the financial aspects of projects. They are based on geolocation and the number of NGA connections in white areas. In case of omission of a number of connections in preparatory stages of the project we end up with underestimating investment costs. In such case during the construction of the network, and while examining the actual situation on-site and in case of covering all eligible users, the situation might result in unplanned increase in investment costs. Due to the high probability of implementation of a combination of technologies, the crucial deviation is the one, recorded in FTTC / FTTH technology. In this case, the increase of absolute terms amounts to approximately 7 140 808 EUR. These costs must then be settled by the operators who are building the network – and from their own resources. On the other hand, in case the construction of acceptable connections was defined through state-of-the art method only in preparatory stages, one will (as already mentioned) not be able to achieve the objectives from DAE because the NGA will cover 31.45% less potential beneficiaries than in case of projects that apply the ODH algorithm.

This study also notes the absence of official records on the number of private households on individual location; it only provides the number of people and the number of private households per individual village/settlement according to Census 2011 (Central Bureau of Statistics, 2011). Since the rules, used for implementation of projects require defining the geolocation for each individual household, the implementation of this procedure is crucial because of the prevailing influence of this number on the natural and financial aspects of projects.

Given that similar research has not been carried out yet, the results, achieved in this research, can be considered as initial results. Taking into consideration previous studies on the effectiveness of the entire cohesion policy, this research is in accordance with findings of Pellegrini, Terribile, Tarola, Muccigrosso, and Busillo (2013) on the necessity of adequate basis for the implementation of co-financed projects.

Considering the justification of a high relevance of a more accurate definition of the number of connections, the developed ODH algorithm considerably increases the level of pairing the locations of potential beneficiaries and unique geolocations of house numbers. With its results, the ODH algorithm contributes to achieving the DAE objectives due to reducing the gap between the input data and the real situation; allowing inclusion of a larger number of eligible potential beneficiaries of the project.

Since research was conducted on a sample of projects that cumulatively cover 42% of the Croatian territory, similar results can be expected in the remaining areas of the country, as well as in less developed countries in the region. This is especially the case because the government records in the region have similar quality characteristics. This matter could therefore be a topic for further research.

Absolutely essential for achieving the DAE goals is therefore the mutual harmonization of national data. The most effective solution is to create centralized records of geolocations of all eligible potential beneficiaries in Republic of Croatia, and in all countries which are prepared to implement projects for the development of broadband infrastructure. Applying this approach, one will avoid differences between projects that refer to the methodology of preparation, while at the same time effectively contribute to achieving the DAE objectives and implement the EU cohesion policy.

#### References

- Bachtler, J., & Turok, I. (2013). *The coherence of EU regional policy: Contrasting perspectives on the Structural Funds*. Routledge.
- Briglauer, W., & Gugler, K. (2013). The deployment and penetration of high-speed fiber networks and services: Why are EU member states lagging behind? *Telecommunications Policy*, *37*(10), 819-835. http://dx.doi.org/10.1016/j.telpol.2013.05.003
- Cambini, C., Briglauer, W., & Grajek, M. (2015). Why is Europe lagging on next generation access networks? *Bruegel Policy Contribution* ISSUE 2015/14, September 2015. Retrieved from http://aei.pitt.edu/68364/1/pc\_2015\_14.pdf
- Carter, D. (2013). Urban regeneration, digital development strategies and the knowledge economy: Manchester case study. *Journal of the Knowledge Economy*, 4(2), 169-189. http://dx.doi.org/10.1007/s13132-012-0086-7
- Central Bureau of Statistics (2011). *Popis stanovništva 2011*. Retrieved from: http://www.dzs.hr/Hrv/censuses/census2011/censuslogo.htm
- Cuaresma, J. C., Doppelhofer, G., & Feldkircher, M. (2014). The determinants of economic growth in European regions. *Regional Studies*, 48(1), 44-67. http://dx.doi.org/10.1080/00343404.2012.678824
- Currie, M., Philip, L. J., & Roberts, A. (2015). Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. *BMC health services research*, *15*(1), 1. http://dx.doi.org/10.1186/s12913-015-0825-0
- da Conceição-Heldt, E., & Meunier, S. (2014). Speaking with a single voice: internal cohesiveness and external effectiveness of the EU in global governance. *Journal of European public policy*, 21(7), 961-979. http://dx.doi.org/10.1080/13501763.2014.913219
- Daly, M. (2012). Paradigms in EU social policy: a critical account of Europe 2020. *Transfer: European Review of Labour and Research*, 18(3), 273-284. http://dx.doi.org/10.1177/1024258912448598
- European Commission (2010b). Europe 2020: A strategy for smart, sustainable and inclusive growth. COM (2010) 2020 final. Retrieved from
  - http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC2020&from=EN
- European Commission (2013). EU Guidelines for the application of State aid rules in relation to rapid deployment of broadband networks, OJ 2013/C 25/01. Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2013:025:0001:0026:en:PDF
- European Commission. (2010a). *A Digital Agenda for Europe. COM (2010) 245 final*. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245R(01)&from=EN.
- Government of the Republic of Croatia (2016). Prijedlog odluke o donošenju Okvirnog nacionalnog programa za razvoj infrastrukture širokopojasnog pristupa u područjima u kojima ne postoji dostatan komercijalni interes za ulaganja. *Documentation from the Croatian Government session*. Retrieved from https://vlada.gov.hr/UserDocsImages//Sjednice/2016/31%20sjednica%20Vlade//31%20-%202b.pdf
- Groot, R. (1997). Spatial data infrastructure (SDI) for sustainable land management. *ITC journal*, *3*(4), 287-294. Retrieved from http://ces.iisc.ernet.in/energy/HC270799/LM/SUSLUP/KeySpeakers/AGroot.pdf
- Gulati, G. J., & Yates, D. J. (2012). Different paths to universal access: The impact of policy and regulation on broadband diffusion in the developed and developing worlds. *Telecommunications Policy*, *36*(9), 749-761. http://dx.doi.org/10.1016/j.telpol.2012.06.013
- Marcut, M. (2014). Social Inclusion and Social Frontiers in the Digital Age-How Can EU's Digital Agenda Help? *Eurolimes*, (17), 53-68.
- McCann, P., & Ortega-Argilés, R. (2015). Smart specialization, regional growth and applications to European Union cohesion policy. *Regional Studies*, 49(8), 1291-1302. http://dx.doi.org/10.1080/00343404.2013.799769
- McCormick, J. (2013). Why Europe matters: the case for the European Union. Palgrave Macmillan.
- Ministry of Maritime Affairs, Transport and Infrastructure of the Republic of Croatia (2014). *Okvirni nacionalni* program za razvoj infrastrukture širokopojasnog pristupa u područjima u kojima ne postoji dostatan komercijalni interes za ulaganja. Retrieved from http://www.mppi.hr/UserDocsImages/MPPI Okvirni program NGA BB%2020-3 14%20FINAL.pdf
- Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T., & Busillo, F. (2013). Measuring the effects of European Regional Policy on economic growth: A regression discontinuity approach. *Papers in Regional Science*, 92(1),

217-233. http://dx.doi.org/10.1111/j.1435-5957.2012.00459.x

- Qiang, C. Z. W., Rossotto, C. M., & Kimura, K. (2009). Economic Impacts of Broadband. Information and Communications for Development 2009: Extending Reach and Increasing Impact, 35-50. Washington, DC. World Bank.
- Rampersad, G., & Troshani, I. (2013). High-speed broadband: assessing its social impact. *Industrial Management & Data Systems*, 113(4), 541-557. http://dx.doi.org/10.1108/02635571311322784
- Reggi, L., & Scicchitano, S. (2014). Are EU regional digital strategies evidence-based? An analysis of the allocation of 2007-13 Structural Funds. *Telecommunications Policy*, 38(5), 530-538. http://dx.doi.org/10.1016/j.telpol.2013.12.007
- Van der Wee, M., Verbrugge, S., Sadowski, B., Driesse, M., & Pickavet, M. (2015). Identifying and quantifying the indirect benefits of broadband networks for e-government and e-business: A bottom-up approach. *Telecommunications Policy*, 39(3), 176-191. http://dx.doi.org/10.1016/j.telpol.2013.12.006
- Villa, P., Molina, R., & Gomarasca, M. A. (2013). Data Harmonisation in the Context of the European Spatial Data Infrastructure: The HUMBOLDT Project Framework and Scenarios. *In Earth Observation of Global Changes* (EOGC) (pp. 179-191). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-642-32714-8\_12

(cc) BY

This work is licensed under a Creative Commons Attribution 3.0 License.