Multilayer, locality aware, telecommunication network deployment algorithm

Submitted: 05/11/2018
Revised: 10/12/2018
Accepted: 02/02/2019

Luis M. Roman[*]
Ante Salcedo[**]
Miguel Alonso Vilchis[***]

Abstract

Purpose – In this paper we propose an iterative approach for the deployment of rural telecommunication networks.

Methodology/approach/design – This approach relies heavily on the concept of locality, prioritizing small 'cells' with a considerable population density, and exploits the natural nesting of the distribution of rural communities, focusing in communities which are populous enough to justify the investment required to provide them with connectivity, and whose sheer size promotes the formation of 'satellite' communities that could be benefited from the initial investment at a marginal expense. For this approach, the concept of 'cells' is paramount, which are constructed iteratively based on the contour of a Voronoi tessellation centered on the community of interest. Once the focal community has been 'connected' with network of the previous layer, the process is repeated with less populous communities at each stage until a coverage threshold has been reached. One of the main contributions of this methodology is that it makes every calculation based on 'street distance' instead of Euclidean, giving a more realistic approximate of the length of the network and hence the amount of the investment. To test our results, we ran our experiments on two segregated communities in one of the most complicated terrains, due to the mountain chains, in the state of Chiapas, Mexico.

Findings – The results suggest that the use of 'street distance' and a local approach leads to the deployment of a remarkably different network than the standard methodology would imply.

Practical implications – The results of this paper might lead to a significant reduction in the costs associated with these kinds of projects and therefore make the democratization of connectivity a reality. In order to make our results reproducible, we make all our code open and publicly available on GitHub.

Keywords: multilayer, locality aware, telecommunication network deployment algorithm.

---

[*] Full-time student of the Masters in computer science at ITAM. E-mail: lromanga@itam.mx.

[**] Full-time professor at the Digital Systems Department at ITAM. E-mail: mvilchis@ciencias.unam.mx.

[***] Full-time student of the Masters in Computer Science at ITAM. E-mail: ante.salcedo@itam.mx. All of them conduct research and consulting on telecommunications industry trends, infrastructure, economics, and public policy.

ROMAN, L. K.; SALCEDO, A.; VILCHIS, M. A. *Multilayer, locality aware, telecommunication network deployment algorithm.* **Law, State and Telecommunications Review**, Brasilia, v. 11, n. 1, p. 145-158, May 2019. DOI: https://doi.org/10.26512/lstr.v11i1.24854

Introduction

Hardly any other industrial sector has had a more relevant role in the furtherance of the technological and economic developments of the last century than the telecommunications sector (Roller, Lars-Hendrik, and Leonard Waverman, 2001). Its relevance relies not solely in its capacity to provide high speed communication between individuals thousands of miles away in regions, hitherto, by all means, inaccessible. But also, in its capacity to automate procedures in a more finned grain fashion than ever due to the advancements in the Internet of Things (IoT) technologies (Wollschlaeger, Martin, Thilo Sauter, and Jürgen Jasperneite, 2017). In fact, it has been shown that the development of such connectivity networks generates positive externalities making the impact of telecommunications infrastructure on economic growth greater than linear. In this same venue, (Röller, Lars-Hendrik, and Leonard Waverman, 2001) found that when levels of infrastructure approach universal service, increasing returns on growth are reached. Along with these developments, it has been noted (Bauer, Johannes M, 2017) that the economic disparity between connected and poorly connected regions has been growing at a staggering rate, leading some to believe that the next differentiating factor between viable and nonviable communities would be the degree of development in this field.

For these reasons, it has become a critical concern for local and federal governments all over the world to provide the legislative and economic groundwork to foster the development of a healthy telecommunications sector. This represents a major challenge for countries where a large percentage of the population lives in rural areas, since, more often than not, the inherent complexity of deploying telecommunication networks over vast inhabited territory and the costs associated with a *design-from-scratch* network constitute an infranqueable entrance barrier.

Since the greatest part of the costs associated with the deployment of a telecommunication network derives directly from the adequacy of its topological design, it is natural that a lot of work in the fields of operations research and network optimization has been focused precisely to this subject. As Prytz (Prytz, Mikael, 2002) pointed out, a good network design is difficult to characterize exactly because it depends on too many factors. A network design engineer typically has to evaluate tradeoffs between several design patterns. Despite the multiple factors involved, Prytz mentions that the main design objectives of telecommunications networks can be summarized as follows:

- Performance: application response time and availability. This frequently translates into requirements on the data transmission rate and the percentage of data that has to be present; or on the probability that an application runs without failures due to congestion.

- Redundancy, resiliency & survivability: network ability to rest available under component failures (like router, switch or network link interruptions). These objectives translate into requirements on network resources, such as the need for alternative paths to reroute traffic.
- Economic viability: the returns on investment (including the cost of equipment, capacity leasing, labor, support, etc.) need to be attractive enough to justify the deployment of the network.

As Prytz shows, there are complex interactions and natural conflicts between the design-objectives explained above; therefore, network designs are frequently broken into smaller subproblems that are easier to deal with. Furthermore, networks are usually subdivided into different segments that obey different functionality and topology, and are often treated separately according to their own architectures. Such segments of a network can be grouped into tiers or layers that are arranged hierarchically. In such manner, the top layer of a network is often referred as the primary, backbone or core, and the subsequent layers known as secondary or tertiary networks. A top-down approach is often used to satisfy each layer’s particular criteria.

Considering that in this paper we are focusing on rural areas we will be assuming that the connected locations are part of an already deployed primary or secondary network whereas, we will be deploying a tertiary network over the non-connected communities. Needless to say, in this context the economic viability factor acquires particular relevance. Thus, in order to address it properly, in this paper we develop a network deployment methodology that highly favors the reutilization of already existing infrastructure and exploits different idiosyncratic notions of proximity that allow the progressive construction of network chunks depending on different coverage and population restrictions, thus, further diminishing the economic hurdle of the project.

The structure of the paper is as follows: First we detail our methodology and its most important implications. Then we go through several experiments that we carried out on two isolated Mexican communities and immediately afterwards we discuss the different results we obtained. Then we go through a thorough examination of previous work on the subject. We present the different nuisances of each approach and how our methodology differs from the principal tendencies along multiple design patterns. Finally, we give some closing remarks and hint some venues for future work.

Methodology

As mentioned in the introduction, our approach aims to make maximum usage of already deployed infrastructure in order to avoid unnecessary expenses. Similarly, it highly encourages the progressive deployment of nested chunks of network, therefore steering clear of excessively large deployments and focusing only on the *connectivity frontier*. With these considerations in mind, the steps of our methodology are the following:

1. Identify a target region to cover. Ideally, it should have a couple of ‘large communities’, large in the sense that their population justifies a network deployment, hence they can be considered as connected centroids for the following iterations. Together with a large amount of ‘small communities. These are going to be the deployment targets.
2. Set a deployment criterion: favor small areas with large populations, these are the ‘low hanging fruits’ or favor large areas with small populations, these are the hard to cover communities and serve as objectives for public policy or to compute the worst-case investment scenario.
3. Once these two criterions have been selected, the algorithm is going to iteratively compute network extensions using Voronoi tessellations centered on the connected localities of the previous iteration. A tessellation partition is selected according to the criterions of point two and a given number of communities is added to the network. The number of communities to connect at each iteration is another parameter of the algorithm. This is going to be carried out until one stop criteria is meat.

    a. Each partition of the tessellation as well as each network extension is constructed using ‘road distance’ in order to give a more down to earth approach of the length of the network, as well as of the investment needed.
    b. It is worth mentioning that at each stage the algorithm is going to favor the integration to the network of communities that are close to roads and whose population is relatively large.

More specifically. Let $C_i{}^k$be the i^th centroid of the k^th iteration, $S^{q_k}{}_{c_i{}^k}$ the tessellation partition associated with it, selected according to criterionq_k and $P(S^{q_k}{}_{c_i{}^k})$ and $N(S^{q_k}{}_{c_i{}^k})$ be the population and number of communities $\pi^t{}_{S^{q_k}{}_{c_i{}^k}}$ t = 1, 2…., $N(S^{q_k}{}_{c_i{}^k})$ in such partition. Then the algorithm goes as follows: For fixed T_p and $T_N$population and number of community’s thresholds, as follows:

1. Let k = 0;
2. Choose $X_0 < N({S^{q_0}}_{c_i^0})$ communities as centroids for the partition according to population criteria (these are the connected networks);
3. Partitionate the communities taking into consideration population and distance to road criteria using the Voronoi tessellation induced by the weighted k-means;
4. While $P({S^{q_k}}_{c_i^k}) > T_p$and $N({S^{q_k}}_{c_i^k}) > T_N$ do;
5. Select ${S^{q_k}}_{c_i^k}$ according to q_k;
6. Add X_k communities ${\pi^t}_{{S^{q_k}}_{c_i^k}}$(where $X_k < N({S^{q_k}}_{c_i^k})$) to the network using road distance and population criteria;
7. Set this X_k communities as centroids for iteration k + 1 and generate a nested partition;
8. Set k = k + 1.

As it can be readily seen, this approach leaves a large set of parameters available to the user, thus, giving her enough flexibility to adapt to situation specific conditions. Furthermore, it guarantees that at each iteration only the nearest, according to road distance, most populous communities are getting connected, therefore, in a greedy approach minimizing expenses and connecting as much people as possible.

Even though our approach is location independent, in the following section we are going to carry out an experiment in one of the hardest regions to connect in Mexico due to its rough geography.

Implementation

In order to test our methodology, we decided to carry out a couple experiments in one of the toughest territories of the Mexican landscape, namely, in two isolated communities in the state of Chiapas called Villa Corzo and Villaflores. This region is particularly hard to cover because most the territory is composed of mountains and dense rainforest. The communities are scantily populated and are widely spread across the fields, hence, providing a worst case scenario for our algorithm. There are two main reasons why we choose to carry out our experiments in Mexico. The first reason is related with the unique legislative landscape the current government has instaured. The second reason is of a more practical nature and it is related with an almost ideally complex distribution of the population. In the following paragraphs we are going to dig deeper into the implications of these motivations and give an outline as to why we consider them exceptionally relevant, then we will give a brief outline of the characteristics of Villa Corzo and Villaflores.

Legislative landscape

In the wake of 2012, the new Mexican government instaurated a set of constitutional reforms regarding several key industrial sectors of the Mexican landscape. Among those legislative interventions, a new set of reglementations was introduced for the telecommunications sector. What was particularly remarkable, was that amid a couple of technical nuisances, the right to connectivity became a constitutional right, therefore, the state committed itself to provide the means, both, in the legislative realm as well as in the infrastructural one in order to guarantee its availability for the gross part of the population, in this case, to the 98% of the population.

Within this context, it has become apparent that the need for efficient ways to provide connectivity to isolated communities has grown into one of the most exacting nature.

Population distribution

As part of its historical heritage, the Mexican landscape has been chiefly characterized by a remarkable tendency to produce large clusters of centralized population. This sole fact has important implications in the design and deployment of a telecommunications network. According to INEGI (INEGI, Total population by locality size and state 2000, 2005, 2010) already by 2017, 79% of the Mexican population lived in urban areas. Nonetheless, as it can be seen in Figure 1, due to the centralized nature of the population distribution this percentage occupies less than 300,000 km^2 which represents less than 10% of the total Mexican territory. Since, according to IFT's stats, nearly 85% of the total Mexican population is already connected, this leaves a mammoth approximate of 300,000 km^2 yet to cover in order to achieve the constitutional objective of 98% of the population. Thus, the need for efficient ways to provide connectivity to isolated communities spread among thousands of kilometers.

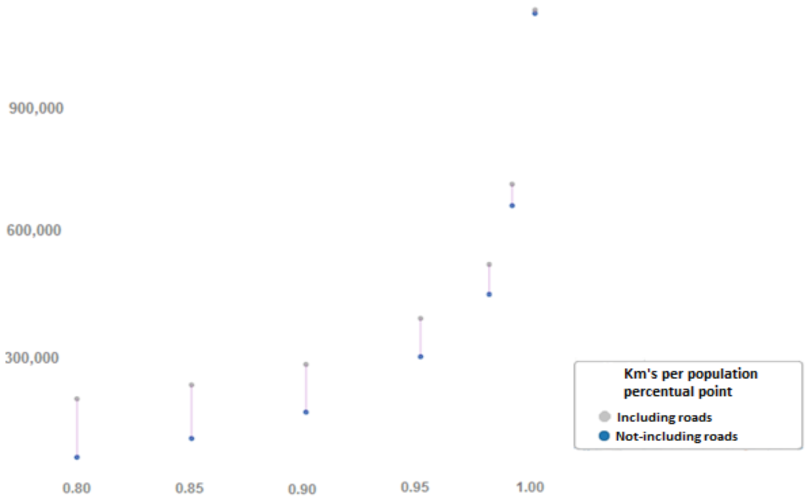

Figure 1 – Kilometers to cover each percentage point of population.

Villa Corzo and Villaflores

The main directives for choosing a region for testing our algorithms were that it must be a poorly connected region, in order to emphasize the field of opportunity, and preferably, it must have a large population, in order to maximize the impact. After searching for the ideal scenario, we came to these two communities, the regions selected in Villaflores and Villa Corzo have a population of 7,109 and 8,794 individuals respectively. As it can be appreciated in Figure 2, these two regions are scantily connected with only available network in the three largest localities. Moreover, the connectivity of these is clearly bad, with most of the sampled points (the sampling was done using OpenSignal's API: NetworkStats) having an unreliable network and an average Arbitrary Strength Unit (ASU) on received signal strength indicator (RSSI) of 19.07761 and an average reliability of 0.13, fitting nicely within the distributional premises of our algorithm.

Additionally, as it can be seen in Figure 3, these two communities have the desirable property that they are well connected with roads and highways, hence, leading to a maximum capitalization of already existing infrastructure, fact that our algorithm exploits splendidly.

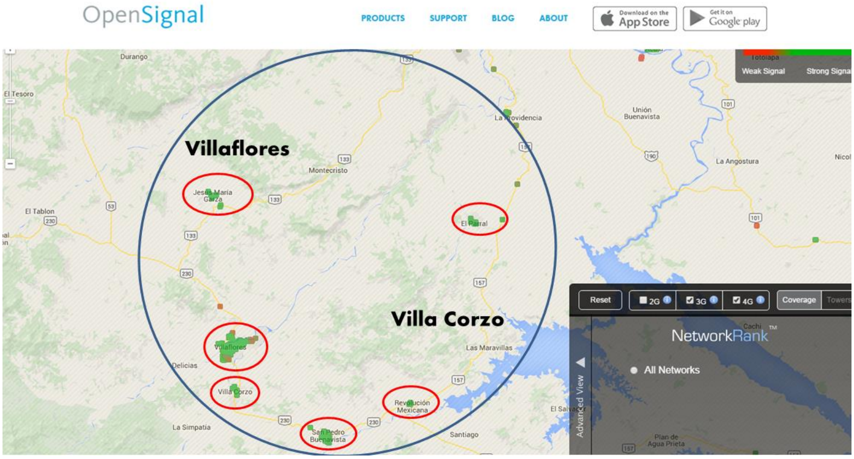

Figure 2 – Communities with broadband access Villa Corzo and Villaflore

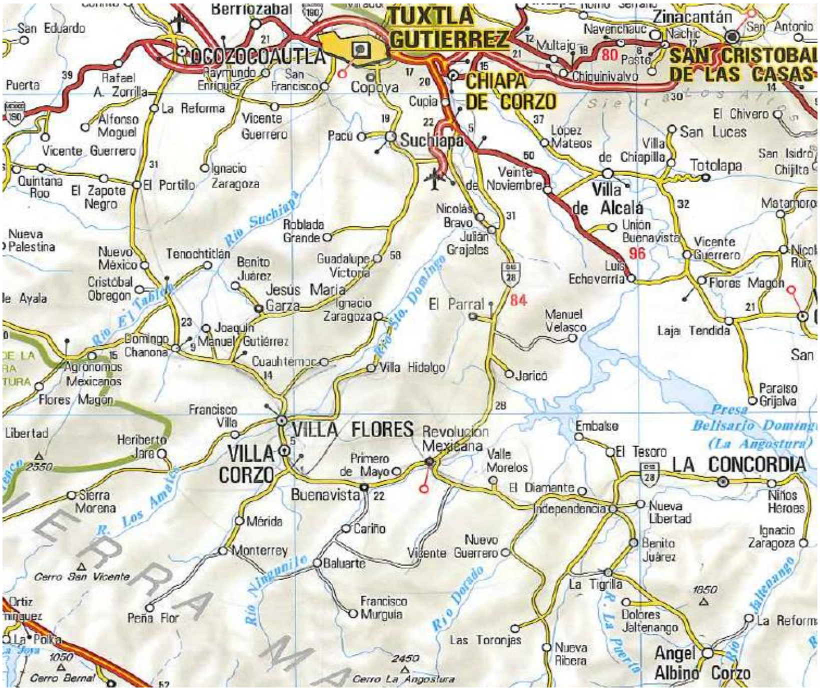

Figure 3 – Roads and trails in Villaflores and Villa Corzo

Experiments and Results

In this section, the main results and findings are shown and discussed. We carefully consider each of the three main steps of our methodology: 1) identify a target region to cover, 2) set a deployment criterion. In this case we will use the stop criteria of minimum population for each cluster, depending on the iteration in which it is. 3) build local networks along roads and trails, and repeat until each cluster meets the criteria of minimum population.

The values chosen as restrictions for the minimum number of citizens in each iteration are: 5000, 750, 100 and 50. These values were the parameters with which we obtained the greatest coverage with the least amount of network deployed.

In the first iteration, the cluster number is adapted to meet the restriction of 5000 citizens for each cluster. To show the methodology, the experiment focused on the region with more population, this as consequence of growth of the localities with tendency to produce large clusters of centralized population.

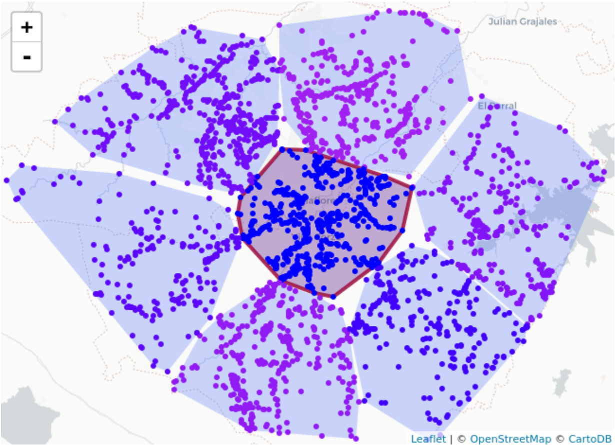

Figure 4 – Selected cluster in Villa Flores and Villa Corzo

In the later iterations, two Google API services were used, the first was Google Maps Roads, this was used to get the distance between a locality and the nearest road (see figure 5).

Figure 5 – Left network topology without distance to road, Right network topology weighted with distance to road.

As it can be appreciated, considering road distance leads to more local shorter networks. Hence it could be a useful tool in case the localities are well connected by roads and trails. Once that value is acquired, we use it as a weight to select the centroids of the partition, in this way, we give a greater weight to the localities with a shorter communication path. We use spanning tree techniques in each region to determine the expansion branches of the secondary network by connecting these locations to the core network along existing roads and trails, this value is available through Google Maps Directions API that returns the real 'road' distance between two locations.

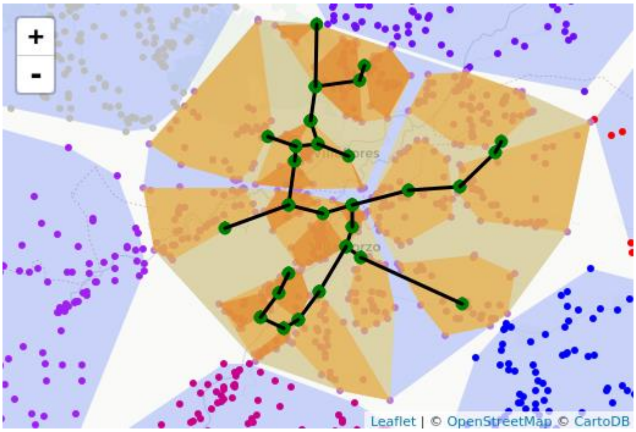

Figure 6 – Spanning tree in Villa Flores and Villa Corzo

To measure the percentage of coverage of the locality with the generator tree, we generate a performance graph, which shows the coverage that each iteration adds, ending with a coverage of 97.6%.

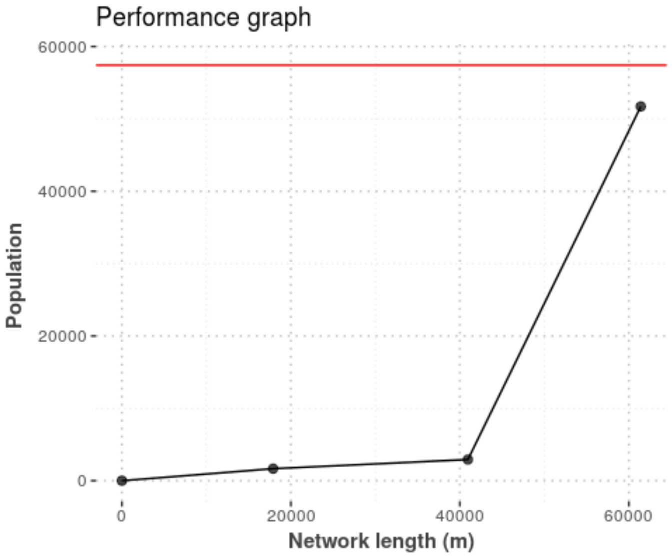

Figure 7 – Performance graph of network coverage

Related Work

As mentioned in the introduction, there exists a huge amount of literature, mainly in the fields of operations research and network theory, that regard the deployment of telecommunications networks as their main concern. It is a well-known fact that the deployment of general capacitated networks is an NP-hard problem, therefore, safe in some unrealistically simplified scenarios, the computation of an optimal deployment is computationally intractable. In order to alleviate this situation, different implementations have suggested multiple heuristics for getting a satisfactory solution, take for example the work of (Ruiz, Efrain, Maria Albareda-Sambola, Elena Fernández, and Mauricio GC Resende, 2015) where a genetic algorithm approach is proposed, they explored multiple encodings as well as parameter configuration in order to acquire acceptable solutions, in this venue, they successfully outperformed some off the shelf implementations, though, the notion of locality and controlled aggregated investment costs are completely disregarded. Our approach is similar in spirit to (Kritikos, M., and G. Ioannou, 2017) in the fact that they use a greedy approach and hence opt for the 'best' decision at each step of the algorithm, whereas regarding the configuration of the nodes we elect a notion of 'proximity' based on the Voronoi tessellation induced by the k-means algorithm similar to (Zhong, Caiming, Mikko Malinen, Duoqian Miao, and Pasi Fränti, 2015).

Conclusion

As a closing remark we believe, as stated by our results, the main contributions of our paper are that it encompasses the different considerations of locality, allows a progressive deployment of the network, encourages the exploitation of already existing infrastructure favoring communities that are relatively 'near' to roads and trails, gives a more down to earth estimation of the costs of deployment since it uses road distances instead of euclidean or large disks distances and finally it is location independent hence its applicability in different scenarios is transparent for the user.

Acknowledgements

This research work has been partially supported by the *Asociación Mexicana de Cultura A.C.*; and by *Centro de Energía y Recursos Renovables del ITAM*.

References

ALBAREDA-SAMBOLA, Maria; FERNÁNDEZ, Elena; RESENSE, Mauricio GC; RUIZ, Efrain. A biased random-key genetic algorithm for the capacitated minimum spanning tree problem. **Computers & Operations Research**, v. 57, p. 95-108, 2015.

ÁVILA, A.; KUHLMANN F.; SALCEDO, A. Searching for economic value in Mexican rural communities. **Proceedings of the 9th Communication Policy Research (CPR) LATAM Conference**, p. 169 -177, 2015.

BAUER, Johannes M. ***The Internet and income inequality: Socio-economic challenges in a hyperconnected society.*** Telecommunications Policy, 2017.

FRANSES, P.H. Fitting a Gompertz Curve. **The Journal of the Operational Research Society**, v. 45, n. 1, p. 109-113, 2014.

FRÄNTI, Pasi; MIAO, Duoqian; MIKKO, Malinen; ZHONG, Caiming. A fast minimum spanning tree algorithm based on K-means. **Information Sciences**, v. 295, p, 1-17, 2015.

INEGI. ***Total population by locality size and state 2000, 2005, 2010.*** INEGI.

IOANNOU, G; KRITIKOS, M. A greedy heuristic for the capacitated minimum spanning tree problem. **Journal of the Operational Research Society**, v. 68, n. 10, p. 1223-123, 2017.

MORA, M.; ROMAN L.; SALCEDO, A. ***Local layered algorithmic model for topological design of rural telecommunications networks.*** International Conference on OR for Development (ICORD), Mexico City, 2016.

OPEN SIGNAL WEBPAGE, ***3G and 4G Coverage Maps***.

ORENDAIN, J. ***Diagnóstico y prospectivas del sector de telecomunicaciones en México, en un entorno de convergencia.*** 2007. Thesis – Instituto Tecnológico Autónomo de México, Mexico City, 2007.

PRYTZ, Mikael. ***On optimization in design of telecommunications networks with multicast and unicast traffic.*** 2002. 36f. Tese de Doutorado – Royal Institute of Technology, Estocolmo, 2002.

ROLLER, Lars-Hendrik; WAVERMAN, Leonard. Telecommunications infrastructure and economic development: A simultaneous approach. **American Economic Review**, v. 91, n. 4, p. 909-923, 2001.

ROMAN, L. K.; SALCEDO, A.; VILCHIS, M. A. *Multilayer, locality aware, telecommunication network deployment algorithm.* **Law, State and Telecommunications Review**, Brasilia, v. 11, n. 1, p. 145-158, May 2019. DOI: https://doi.org/10.26512/lstr.v11i1.24854