Using self organizing neural networks to improve optical network design

Alexandru Smeureanu, Stefan D. Dumitrescu, Liviu A. Cotfas, Andreea Diosteanu

***Abstract*—**The paper presents a model for data (IP) communication network design that uses self organizing neural networks and graph theory adapted algorithms to improve its reliability and efficiency. The resulting architecture is a layered hierarchical network; the first layer represents the core of the network and has a ring topology while the second layer has a tree topology with a maximum depth constraint. The model is aimed towards providing a close to optimal solution for large problem configurations that are unsolvable in polynomial time.

***Keywords*—**self organizing neural network, graph algorithms, graph models

I. Introduction

THE paper presents the development of an optimum optical infrastructure planning model based on self organizing neural networks. This system has been added as a functionality into an preexisting Network Management System[1].

The strong demand for Internet services in the recent years has had a great impact in the development of communication networks. Because rapid growth was, and still is, based on ad-hoc decisions taken in the absence of a development plan, in time it has led to rather poor network architectures (incurring efficiency and reliability issues) and other inherited design problems. Because of ineffective management and lack of understanding in regard to the expansion or assimilation process, in almost all situations technical departments are still subjected to very strict budgets and time pressures. This forced them to make compromises and quickly incorporate parts of the newly acquired infrastructure. This in turn leads to suboptimal designs, loss of quality, inheritance of existing network and design problems and even lack of complete information or control.

Another cause leading to an uneven expansion of a network is related to the non-uniform distribution of internet service demand. Over the last decade the Internet has evolved very rapidly from a relative scarce commodity present only in universities or as a luxury item (available only to a rather wealthy segments of population) to an affordable, common-place service (following a drastic decrease internet connection pricing) and even to a state guarantee right in countries like Finland. The geographical areas needed to be covered have started from just several interest points of wealthy districts to cover to a high degree entire countries.

The European Union has stated its support for all programs identified as Information Society [13] which have become a key priority for structural funds interventions with a strong emphasis on demand for services and applications. These programs are designed to convince and support ISPs to invest and develop new infrastructure in areas that will otherwise be deemed unprofitable. This strategy opened the race for Internet infrastructure development in rural areas.

Another usual form of company development is to buy-out other smaller, local companies. Because of high acquisition costs, most often companies do not afford to invest money in the assimilation process. Normally during the assimilation process the best approach would be to keep the infrastructures separate, extend the existing well-thought infrastructure, and migrate the customers to it and the end shutdown of the unused old infrastructure. This process can be very expensive especially if the area of operation of the two internet service providers is not overlapping.

The current trend of technology in the last decades has shown optical networks to be the preferred choice in infrastructure development both for the core of the Internet and last mile connections in metropolitan area networks. The advantages of using fiber optics in metropolitan area networks is that it has virtually no constrain on cable length. On the other side, active terminator equipments are expensive.

The introduction of Voice and Video over IP (VoIP and VVoIP) communications and IPTV service has increased the need for reliable and robust network designs. One of the best ways to increase the reliability and robustness of a network is

This work was supported in part by the following projects: 1. „Doctoral Program and PhD Students in the education research and innovation triangle”. This project is co funded by European Social Fund through The Sectorial Operational Programme for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies (Project no. 7832, “Doctoral Program and PhD Students in the education research and innovation triangle, DOC-ECI”). 2. POSDRU, financing contract POSDRU/6/1.5/S/19

Alexandru Smeureanu is a graduate of the Academy of Economic Studies in Bucharest, Romania as well as of Politehnica University of Bucharest. He is currently undergoing a PhD in the Academy of Economic Studies. (e-mail: alexandru.smeureanu@gmail.com)

Stefan D. Dumitrescu graduated from, and is currently a PhD student at the Politehnica University of Bucharest, Romania. (e-mail: dumitrescu.stefan@gmail.com)

Liviu A. Cotfas graduated from, and is currently a PhD student at the Academy of Economic Studies, Bucharest, Romania. (e-mail: lcotfas@gmail.com)

Andreea Diosteanu graduated from, and is currently a PhD student at the Academy of Economic Studies, Bucharest, Romania. (e-mail:andreea.diosteanu@gmail.com)

to introduce redundancy.

Redundancy can be achieved in several ways in optical networks.

II. Redundancy in IP networks

The best way to ensure redundancy in IP networks is to have two different physical (and geographical) optical circuits between each node. In this situation, because the two links start in the same device and end in another, they can be aggregated at Layer 2 level using a link aggregation or port trunking technologies such as EtherChannel [2]. The advantage in this case is that the two links can be used simultaneous and bandwidth is cumulated. If one of the links fails there is very little down time (less than 1 second). However this solution is not cost effective because is doubles both the number of links and interfaces needed for every device in the network.

The less expensive solution is to use ring topologies together with Layer 2 protocols or Layer 3 protocols to do load balancing and reroute traffic when one link fails. However, these protocols have higher convergence time in case of a topology problem and the network is still vulnerable if two distinct incidents occur in the same time. This will be reflected further on.

Giving the economical constraints of the technical solutions, the Internet Service Providers mainly rely on ring topologies and use optical cables that enhance reliability and offer high bandwidth capabilities. In order to improve redundancy, the network is split into several regions that have important devices interconnected in a ring.

The goal of our model is to obtain a near-optimal solution to connect several locations within a city, respecting a given network hierarchy and minimizing the cost expressed as geographical distance (i.e. minimizing cable length).

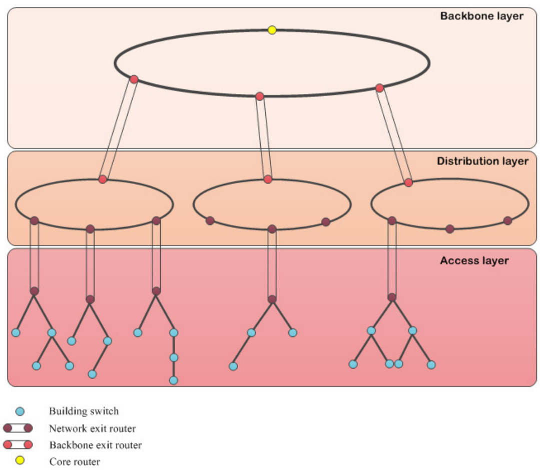

Fig. 1 Multi-layer network architecture

Giving the economical constraints of the technical solutions, the ISPs mainly rely on ring topologies and use optical cables that enhance reliability and offer high bandwidth capabilities. In order to improve redundancy, the network is split into several regions that have important devices interconnected in a ring.

The resulted network topology is a multi layer hierarchical network (fig 1).

- *The first layer is the backbone layer.* The backbone layer is a fiber optical ring that uses high bandwidth optical interfaces and covers a large area in a city. This layer is designed to offer a high capacity path for the aggregated traffic to exit the ISP network. There are two types of routers operating on this ring. The Backbone Exit Routers (BER) and the Core Routers (CR). BER routers are the communication gateway between the distribution layer and the backbone layer and in consequence are part of both layers. The CR is gateway with the exterior of the ISP network meaning that all the traffic leaving the network if forwarded outside by it.
- *The second layer is the distribution layer.* This layer is formed out of several independent fiber optical rings. Each of these rings is connected to the backbone layer via a BER router and contains multiple network exit routers (NER). NER routers are the gateway between the distribution layer and the access layer and belong to both layers.
- *The third layer is the access layer.* This layer is formed out of several parallel access sub networks that interconnect costumer locations with the distribution layer. The access sub networks have a hierarchical topology starting with the root node being a NER.

The presented network topology has a high degree of resilience and is cost effective because it minimizes the use of high bandwidth, long distance optical transmitters. The topology is also in accordance with the Cisco recommendations for building scalable network further more this topology is used by several large ISPs operating Europe such as Swisscom and Airbites.

III. Problem statement

Given a set *V* containing geographical locations that need to be interconnected in *n* distribution rings and m_i access sub-networks. It is requested to generate an optimum graph *G(V,E)* representing the network infrastructure, where *E* represents a set of links between locations in *V,* and *G(V,E)* has the following properties:

- $\exists K$ where $K \epsilon V$ and *K* is core router
- $\exists G'(V', E')$ which is a sub graph of *G(V,E)* where |*V'*| = n+1 (| | cardinal operator) and |*E'*| = n+2 and *E'* forms the Hamiltonian circuit and $K \epsilon V'$ (the backbone ring)
- $\exists G''_i(V''_i, E''_i)$ with $0 \leq i < n$ a set of sub graphs of *G(V,E)* where |*V*| = n+1 (| | cardinal operator) and |*E'*| = n+2 and *E'* forms the Hamiltonian circuit (the distribution rings) and
    - $V''_i \cap V''_j = \emptyset$ where $i \neq j$, $0 \leq i, j < n$

- $V_i'' \cap V' = \{(x_i, y_i)\}$ where $0 \le i < n$ and $d\left((x_i, y_i), \left(\overline{x_{V_i''}}, \overline{y_{V_i''}}\right)\right)$ (the BER)

- $\exists G_{i,j}'''(V_{i,j}''', E_{i,j}''')$ where $0 \le i < n$, $0 \le j < m_i$ and:
  - $V_{i,j}''' \cap V_{m,n}''' = \emptyset$ where $i \ne j$, $0 \le i, j < n$
  - $|V_i'' \cap V_{i,j}'''| = 1$ where $0 \le i < n$

- Each $G_{i,j}'''(V_{i,j}''', E_{i,j}''')$ is a minimum spanning tree of the complete graph formed with the $V_{i,j}'''$ where $0 \le i < n$, , $0 \le j < m_i$ and n << |V| (| | cardinal operator)

- $V_{i,j}''' \cap V_i'' = \{(x_{i,j}, y_{i,j})\}$ where $d\left((x_{i,j}, y_{i,j}), \left(\overline{x_{V_i''}}, \overline{y_{V_i''}}\right)\right)$ is minimal. (the NER is the point closest to the geographical center of the sub network)

- The sum of all the Euclidian distances of the links (edges) in E is close to minimal.

The problem is focused at this point on the economical constraints, the robustness of the network is ensured by the network architecture. The resulted network will have a topology similar to the one presented in Fig. 1

The problem complexity is easy to assess by decomposition of the problem in three different sub-problems equivalent to each layer of the network. At the first layer, the backbone layer:

- the first step is grouping the locations in clusters further referenced as "regions" based on their geographical connections. The complexity and quality of the grouping is strictly related to the clustering algorithm and for the moment we can assume it to be polynomial. The locations closest to the cluster centroids will designated to be BER locations;
- the second step is to solve a Hamiltonian circuit, also known as the Traveling Salesman Problem (TSP) [3] for the nodes selected to be part of the backbone ring (BER). The formulation of the TSP problem is: find the minimum-length tour trough N points (or 'cities') by visiting each point only once. Although the problem seems simple and humans are known to offer good solutions even for medium size ones [4], the problem is NP-complete [5] and thus unsolvable in polynomial time at the moment. Overall, the algorithm complexity assuming that the clustering algorithm is ideal and offers the optimum clustering solution, is still going to be NP-Complete because of the TSP problem [8].

At the second layer, the distribution layer:

- the first step is to distribute a number of sub networks that will be created in each region. This can be done using a repartition method based on the number of location present in the region. However a better method is to distribute the sub-networks based on previous traffic demand or subscription type, if that information is available. Each region will have m_i sub networks;
- the second step is further grouping the locations in the current region in clusters that will make up the access networks. The clustering is done again based on geographical locations and information about traffic demand. The number of clusters in region is equal to m_i. The nodes closest to the cluster centroids become the NER;
- the third step is to link all the NER in one region in the optimum ring and to solve the Hamiltionian circuit problem.

At the third layer, the access layer:

- the first step is to generate all the minim spanning trees (MST) for each of the clusters starting with NER as the centers of the network.
- the second step is to run a heuristic algorithm to reoptimize the network structure to improve robustness by reducing the maximum depth of the tree and reduce the risk of overloading one link by ensuring that nodes that consume more traffic are linked closer to the root node. The network resulted after the re-optimization process is not necessarily a MST solution anymore. However the process can be controlled by using the minimum cost (a MST solution that provides by default the minimum cost) as a reference and making sure that the new network doesn't deviate more than a certain percent from it.

The algorithm uses two steps that involve finding the smallest ring that will pass through all the locations and have the minimum cost. As seen in step two for the first layer the complexity of solving this problem is NP-complete making the entire algorithm NP-complete. Since exploring the entire solution space is not possible because of the high number of possibilities, approximations algorithms are used.

The algorithm used to solve the optimum ring problem is based on self organizing neural networks.

IV. Problem Solution Using Self Organizing Neural Networks

A. *Self organizing neural network*

Self organizing neural networks are categorized as unsupervised neural networks. They are designed to discover correlations in the input data and to form categories. This type of neural network is inspired by the property of the human brain to self organize upon receiving inputs presenting similar patterns.

In literature neural networks and especially associative neural networks have been used to solve optimum problems [6]. The general approach in these problems as well as in our situation is to define an energy function similar to the potential energy in physics.

The stability of the entire system is determined by the energy function. Similar to physics, the lower the energy of the system, the more stable it is. All the imposed restrictions of the optimum problem need to be reflected in this energy function. The tendency of the system to shift towards a lower energy configuration during the simulation is the only mechanism leading to a solution for the optimization problem.

In this case this method is applied to solve two combinatory problems for the minimization of distance between centroids and the nodes from the same cluster (the clustering problem) and between nodes that make up the same ring (the Hamiltionian circuit).

Depending on the problem addressed the self organizing neural networks have different topologies that incorporate the restrictions of the problem.

B. Solving the partitioning problem

The general cluster analysis is composed of two important steps: determination of the number of clusters needed and the classification of the data into clusters. In the proposed model the cluster analysis focuses on the second part. A self organizing neural network algorithm is used to group the locations in geographical regions that later are interconnected to form a given number of sub-networks which is an input parameter of the model.

The self organizing neural network consists of an input layer and the Kohonen layer. The Kohonen layer is usually designed as a two-dimensional arrangement of neurons that maps an N-dimensional input to two dimensions, preserving topological order, but for the purpose of identifying cluster membership, we use a one-dimensional Kohonen layer.

The input layer of neurons is fully connected to the Kohonen layer. The Kohonen layer computes the Euclidean distance between the weight vector or more intuitively the current location of the neurons in the 2D space and the input pattern. This is done for all the neurons in the network. The closest neuron in terms of Euclidian distance is selected winner and is activated together with the other neighbor neurons.

The Kohonen layer of our model is a one-dimensional array of neurons, indexed by i where $0 \le i < n$. The index indentifies a specific location and also indicates the neighbor relation between two nodes by the absolute difference between the two indexes. Because self organizing neural networks are trained by an unsupervised competitive learning algorithm, a process of self organization takes place [7].

The neuron that was declared winner updates its weights and migrates towards the current input location. The migrated distance is equal to the distance between the neuron and the current location multiplied by the learning rate. Similar to the winner neuron, the neighbor neurons are also updated with the same value multiplied by a penalty function also known as the neighborhood function.

$$d(i,w) = |i - w|$$

d absolute index distance function
i the coordinates of current location
W the winner neuron index

$$f(d,\theta) = \begin{cases} e^{-(d\theta)^2} & where\ d < n/10 \\ 0 & \end{cases}$$

$f(*,*)$ the neighborhood function
n the total number of sub networks
θ neighbor influence loss coefficient

In one epoch all the locations are used as input of the neural network once. The network will enter another epoch if at least one neuron has changed its location in the previous epoch with a distance $\Delta d > 0.1$. This is also known as the convergence condition.

During the self organization process, the weights of the neurons asymptotically converge towards the values of cluster centroids. These weights are used afterwards to partition the locations into geographical regions. All the locations from one region will be interconnected by the same sub-network. When the neural network has converged, the weights of each neuron represent the coordinates of a cluster centroid. Each location is then associated with the cluster whose centroid is the closest.

The clustering algorithm is the following:

1. Initialize n neurons with random position distributed around the geometrical center of the locations $(\overline{x_V}, \overline{y_V})$. Initialize learning rate $\mu = 0.6$, energy loss coefficient $\theta = 0.1$, epoch counter $i = 0$ and $\alpha = 0.12$
2. Initialize a list that contains all the locations V and perform a random shuffle
3. Initialize a list with the initial location of each neuron
4. For each location in the list
   - For each neuron select the neuron w closest to the current location
   - Adjust the neighbor neurons keeping in mind the index distance from the current neuron in the hierarchy
5. Increment epoch $i = i + 1$
6. Adjust the neighbor influence loss coefficient by multiplying it with the learning rate. $\theta_i = \mu * \theta_{i+1}$
7. For each neuron calculate the distance Δd covered during the last epoch using the list saved in step 3.
   - If $\Delta d < 0.1$ go to step 2.
8. For each location in V find the closest neuron and assign the location to the cluster represented by the neuron

C. Solving the optimum ring topology

In the adaptive neural network approach, a set of nodes connected by a certain type of interconnected network is modeled. Based on the inputs that a node receives, a certain output is computed and the results are propagated to the other nodes. The strength of the output propagation signal in the network is dictated by the self organizing mechanism. The major difference between the Hopfield neural networks which are not trained based on the problem data is that the adaptive learning neural networks are trained by repetitive presentation of input data.

The heavy interconnected neural networks characteristic to Hopfield networks are replaced with a simple neural network topology with $O(n)$ nodes and $O(n^2)$ connections. In terms of topology, the self organizing neural networks are also similar to elastic networks. The difference between the two network types is the adaptive training mechanism.

Both elastic networks and self organizing networks start from a ring shape topology that is gradually elongated towards each node ('city'). When the network converges the ring will represent the tour. The energy function used in self organizing networks is based on a competitive mechanism designed to stretch the ring towards each city.

Kohonen advises that the convergence speed is higher if the initial neighborhood function *f(x)* is rather wide in the beginning allowing fast competitive learning and is gradually decreased as epochs pass. The neighborhood function plays an important role in the determination of good quality solutions and also the time need for the network to stabilize. Its main characteristics are:

- In the beginning to assign a larger importance to the nodes near the winner node so that these nodes move faster tot the location presented as the input set
- This attraction force should decrease over time as epoch pass

In literature there are several such neighborhood functions coupled with other modifications: the *Angeniol method* permits the alteration of the neural network structure. The number of neurons increases according to the creation mechanism which is essential for obtaining near-optimal results. In the neuron creation process, a neuron will be duplicated if during one epoch it has been chosen as the winner for two different locations.

The neural network structure will be altered and the newly created neuron is inserted into the ring as a direct neighbor of the winner node. Both the new added neuron the winner neuron are prohibited from being the declared winners in the next competition epoch.

This approach guarantees that the identical neurons are separated because of the moves of their neighbor neurons. In addition a deletion process occurs if a neuron has not been chosen as the winner during three complete cycles of iterations. The neighborhood function used in the Angeniol method is:

$$f(Gn) = \frac{1}{\sqrt{2}} e^{\frac{-n^2}{G^2}}$$

n the cardinal distance from the current neuron to the winner

The *Guilty net* method is based on a competitive learning network with a conscience mechanism which is employed to separate neurons that have converged to the same city. One of the pitfalls of the Kohonen algorithm is its inability to update all of the nodes properly, allowing neurons to get stuck out in isolated regions from where they never get out, while some other neurons learn many times and start to oscillate between two locations. The conscience mechanism proposed by Desieno is based on the simple idea to produce equiprobable weights in Kohonen learning. The method restrains the neurons from claiming several locations in the same round. The proposed algorithm for solving the optimum ring topology uses a neural network in which the neurons number is twice the number of locations that need to be interconnected.

Other hybrid methods that include genetic algorithms combined with neural network [15] have also been considered, however their results have not been mentioned for this problem.

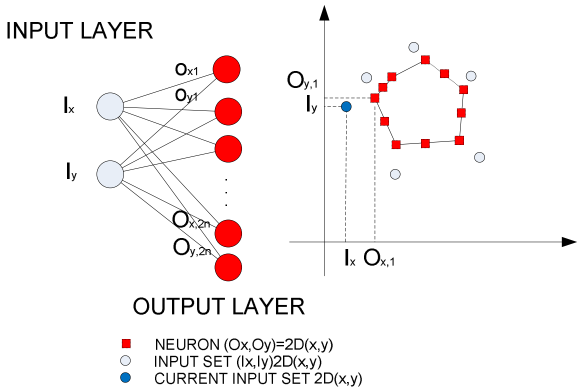

Fig. 2 Conventional neural network view and 2D view of neurons' positions

The inputs of the neural network are two-dimensional and represent the coordinates of the *n* location points. There are exactly *2n* neurons. Each has two weights corresponding to the two inputs that also represent the output. In addition to the traditional neural networks, the neurons are linked together in a ring topology. Each neuron *i* is linked with two adjacent neurons with id ((i+1) mod n) and ((i-1+n) mod n).

In the initial phase the neurons are initialized and located on a ring that is centered around the center of the distribution. In the current model, the center of the ring is placed around the geometrical center of all the BER locations and for each sub-network around the geometrical center of the NER responsible for that sub-network.

After placement, an iterative process takes place: in each epoch a complete list of locations is generated and randomly shuffled. For each location, the neuron closest to the location is declared the winner. A neuron can only be a winner once in epoch and consequently only for one location. The weights of the neuron are adjusted as to shift it towards the desired location. The shift is equal with the difference between the winner neurons coordinates and the location coordinates, multiplied by the learning rate. The rest of the neurons are updated similarly by the formula:

$$O_{i,t+1} = O_{i,t} + \mu * f\big(d(i,w)\big) * (I_c - O_{i,t})$$

$O_{i,t}$	the output of neuron *i* in epoch *t*
I_c	the coordinates of current location
μ	the learning rate
w	the winner neuron index
$f(x)$	the neighborhood function
$d(i,w)$	absolute index distance function

$$d(i,w) = \min\left(n - |i - w|, |i - w|\right)$$

n	the total number of locations
i	the coordinates of current location
w	the winner neuron index

$$f(d,\theta) - \begin{cases} e^{-(d\theta)^2} \; where \; d < n/5 \\ 0 \end{cases}$$

n	the total number of locations
d	absolute index distance function
θ	energy loss coefficient

The iterative process is repeated until stabilization is achieved. Stabilization is achieved when all locations have been occupied by a neuron.

The algorithm for solving the optimum ring topology is:

○ Initialize *2n* neurons with the coordinate of a circle located in the geometrical center of all the locations $(\overline{x_V}, \overline{y_V})$. Initialize learning rate $\mu = 0.6$, energy loss coefficient $\theta = 0.1$, epoch counter $i = 0$ and $\alpha = 0.12$

○ Initialize a list that contains all the locations designated to be part of the core ring and random shuffle the list.

○ Initialize neuron winner list *lw*

○ For each location in the list

- For each neuron select the neuron *w* closest to the current location that is not in the list *lw*
- Adjust the neuron's weights and add it to the *lw* list
- Adjust the neighbor neurons keeping in mind the distance from the current neuron in the hierarchy

○ Adjust the energy loss coefficient by multiplying it with the learning rate.

○ $O_i - \mu * O_{i+1}$

○ Increment epoch

○ $i = i + 1$

○ For each location *L* in the list

○ Initialize S = 0

○ For each neuron select the neuron closest to the current location as P

○ $S = S + d(l., P)$

○ If $S > 10^{-8}$ go to step 2.

○ For each neuron in index order *i* find the location L_i

○ Skip if is the first one

○ Add to *E* part of the problem solution *G(E,V)* the link (L_i, L_{i+1})

This proposed algorithm has shown satisfactory capability for providing high quality results with reasonable computational efforts. In contrast to other approaches, the algorithm used is easy to configure and adjust. A black box test using a subset containing large and medium problems has shown that the algorithm provides solutions with results within 1.3% of optimality in a fraction of the time required to determine the optimal solution, as will be further detailed.

V. Model Performance and Results

In order to test the performance of the model, two standard problem scenarios have been used.

- The first scenario is defined by 200 locations that need to be connected along with their geographical locations and 10 levels maximal depth allowed in a sub-network. All the points are geographically located within an area of 2.5 km x 2.5 km.
- The second for scenario is defined by 2000 locations that need to be connected along with their geographical locations and 10 levels maximal depth allowed in a sub-network. All the points are located within a 20 km x 20 km area.

The performance of the solution is evaluated based on the cost of implementation as a function of the total amount of fiber optical cable used (less is better) along with the robustness of the network (more is better).

In order to evaluate the robustness, several network failure events are randomly generated and used in all test cases in order to preserve consistency. All network events have the same duration in time, are not concurrent and affect only one node at a time. However, because of the network topology, a failure can affect more than one node. The first node of the ring is considered to be the network exit router. If during a network failure a node has no path towards the exit router, it is considered to be down. The robustness of the network is measured as the number of operational nodes divided by the number of network events and divided by the number of nodes.

$$A(x) = \begin{cases} 1, x \; connected \; to \; the \; exit \; node \\ 0, if \; not \end{cases}$$

A(*)	the availability function
x	the current node

$$I_r = \frac{\sum_{i=1}^{NE} \sum_{j=1}^{N} A(j)}{NE * N}$$

I_r	the network robustness indicator

NE	the total number of events
N	the total number of nodes

The results for the first scenario are:

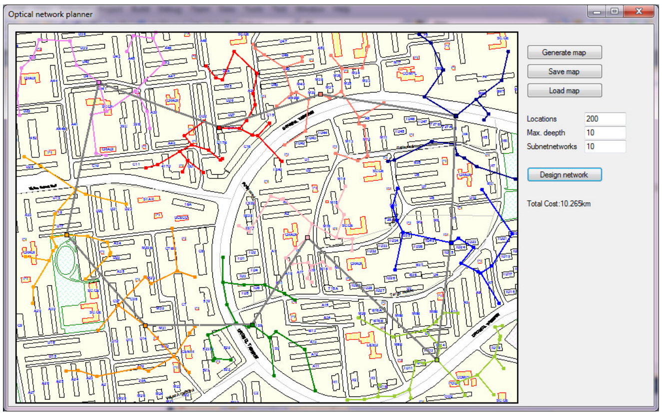

Fig. 3 Model proposed network structure

The solution proposed by the model can be seen in fig.3. Because of the rather low number of points, only one region has been created.

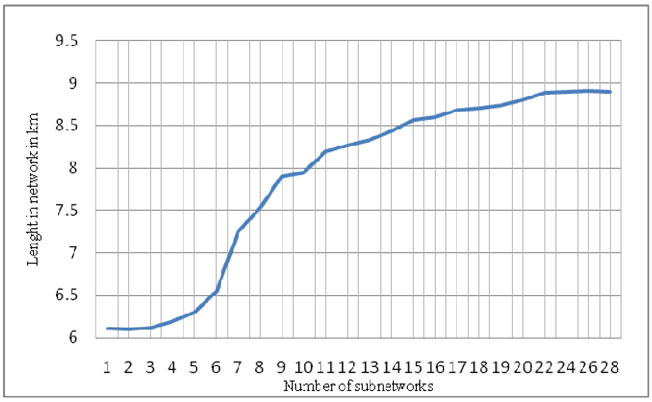

Fig. 4 Total cable length of the proposed infrastructure

The chart in fig. 4 reveals the relationship between the total amount of fiber optical cable needed to build the proposed the infrastructure and the number of sub-networks.

The results indicate that the total amount of cable increases with the number of sub-networks. Also, it shows that there are two asymptotic limits both of the minimum amount and at the maximum amount of cable needed. These limits correspond to less than three sub-networks and respectively more than twenty sub-networks.

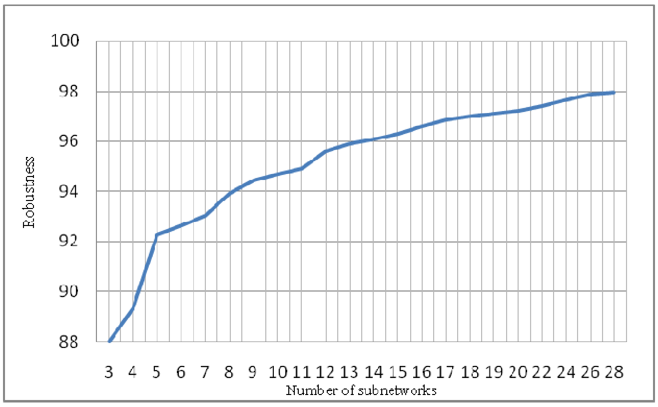

Fig. 5 Network robustness as a function of number of sub-networks

The result clearly shows the increase in robustness as the number of sub-networks increases. However, the increase is not linear, but more logarithmic in nature, as the robustness increase amount is getting smaller as more sub-networks are added. It is interesting to note, that even for three sub-networks linked in a triangle shape, the robustness is rather high (88%).

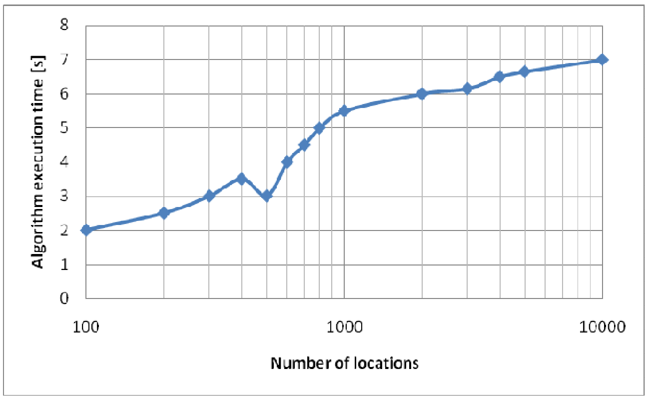

Fig. 6 Algorithm execution time as a function of number of locations

By testing the performance of the algorithm as a function of number of locations the complexity of the neural network clustering algorithm and the sub network generator is evaluated.

The chart in fig. 6 presents the performance of the algorithm when the model is ran for different size networks. On the X axis the number of locations that need to be interconnected is represented on a logarithmic scale. On the Y- axis the algorithm running time in seconds is calculated. These measurements have carried on a problem containing a constant number of sub networks equal in a network spitted in three regions. The results show that algorithm has a logarithmical behavior in respect to the number of locations.

The important results are not the absolute time needed to solve the problem, that can differ based on the implementation, but rather the dependency between time needed by the algorithm and number of locations which is logarithmical.

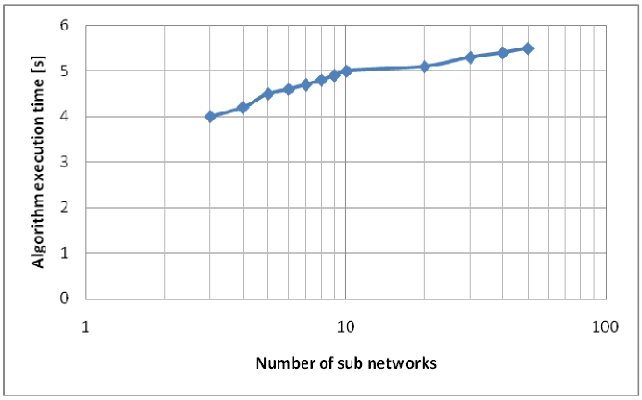

Fig. 7 Algorithm execution time as a function of number sub networks

By testing the performance of the algorithm as a function of number of sub networks the complexity of the optimum ring topology performance can be analyzed.

The chart in fig. 7 presents the performance of the model as a function of number of sub-networks. The X-axis contains the number of sub-networks in a logarithmic scale. Similar to the previous chart, the Y-axis presents the running time of the algorithm when all other components are kept contestant. The total number of locations that need to be interconnected is in this case 800 while the total number of regions is three. The tests have included a varying number of sub-networks of 3 to 50.

The results show that the algorithm execution time follows a linear shape distribution on a logarithmical scale. This indicates that the algorithm has a logarithmical complexity in the number of sub networks detected.

VI. Conclusions

The purpose of our work was to create a network design model that allows for efficient planning of a large area data network. The model has been calibrated and integrated in a network management system [9] that works with real geographical data.

The proposed network design is a balance between cost and redundancy. Because of the hierarchical structure of the network solution and the methods involved in solving the problem, it is impossible to obtain an exact optimum solution. However, the self organizing neural network algorithms used offer a very close approximation of the optimum solution. In [10] solutions for solving the optimum ring topology show results within 1.3% of optimality in comparison to other traditional approximation methods that are within 3% of optimality.

The model can also be used by management as a method of assessing the cabling cost involved in building the infrastructure.

Although in real applications is not always possible to implement the exact proposed network and adjustments need to be made, the proposed network layout still offers very good results or at the least an almost optimum starting point. An acceptable error margin can be calculated and added to the model similar to the one technical teams take in account while on-field. These margins usually depend on the topology of the environment (ex. urban, rural, aerial cable, underground cables) and will need to be adjusted accordingly.

Further extension of the model will take in account elements such as: the cost of the network, the price of the media convertors, router interfaces and operational costs related to equipments.

Other concepts such as priority mechanisms in multi operator Fiber-To-The-Building access network concepts [12] and vulnerabilities assessment and security similar to the ones presented in [14] will also need to be take into account in the future research.

For a more realistic modeling of the sub network, information from the GIS system could be used, such as: road, underground ducts and building infrastructure data.

References

[1] Dumitrescu S., Smeureanu A., Diosteanu A., Cotfas L. A., *Adaptable Network Management System using GIS and network ontology*, 9th RoEduNet International Conference, 2010, pp.310–315.

[2] Cisco online white paper, Cisco EtherChannel Technology http://www.cisco.com/en/US/tech/tk389/tk213/technologies_white_paper09186a0080092944.shtml#wp34608, retrieved, Jan 2011

[3] Gutin G., Punnen A. P., *The Traveling Salesman Problem and Its Variations*, Springer, 2006.

[4] MacGregor J. N., Ormerod T., *"Human performance on the traveling salesman problem"*, Perception and Psychophysics Vol.4, No.58, pp. 527–539

[5] Garey MR, Johnson DS, *Computers and Interactability:A Guideto the Theory of NP-completeness*, Freeman: San Francisco, 1979

[6] Hopfield JJ,Tank DW, *Neural computation of decisions in optimization problems.* Biological Cybernetics No.52,1985, pp141-152.

[7] Burke LI., Ignizio JP. , *Neural networks and operations Research*, Comp Ops Res, Vol.19, No.4, 1992, pp. 179-189.

[8] Cook, William; Espinoza, Daniel; Goycoolea, Marcos, *Computing with domino-parity inequalities for the TSP*, INFORMS Journal on Computing, Vol.19, No.3,2007, pp 356–365.

[9] Smeureanu A., Dumitrescu D.S., *Implementing a GIS application for Network Management*, Journal of applied quantitative methods, Bucharest, Vol.5, No.1, 2010

[10] Somhom, A. Modares, T. Enkawa, *A Self-Organising Model for the Travelling Salesman Problem*, The Journal of the Operational Research Society, Vol. 48, No. 9 (Sep., 1997), pp. 919-928

[11] Sagy Bara, Mira Gonena, Avishai Wool, *A geographic directed preferential internet topology model*, Computer Networks, Vol.51, No.14, 2007, pp.4174-4188

[12] Germán Santos-Boada, Jordi Domingo-Pascual, *Priority Mechanism for Multioperator FTTB Access Networks*, 12th WSEAS International Conference on COMMUNICATIONS, Heraklion, Greece, July 23-25, 2008

[13] G. Liodakis, I.O. Vardiambasis, V. Zacharopoulos, And I. Kaliakatsos, *Teleworking through Optical Networks, Wireless Links, and PBX Technologies,* Proceeding TELE-INFO'05 Proceedings of the 4th WSEAS International Conference on Telecommunications and Informatics, Stevens Point, Wisconsin, USA, 2005

[14] Stamatios V. Kartalopoulos, Di Jin, *Vulnerability Assessment and Security of Scalable and Bandwidth Elastic Next Generation PONs*, Proceedings of the 11th WSEAS International Conference on COMMUNICATIONS, Agios Nikolaos, Crete Island, Greece, July 26-28, 2007

[15] Jiri Stastny, Vladislav Skorpil, Genetic Algorithm and Neural Network, Proceedings of the 7th WSEAS International Conference on Applied Informatics and Communications, Athens, Greece, August 24-26, 2007