

# Graph Partitioning for Survivability in Multi-Domain Optical Networks

Hamza Drid, Bernard Cousin  
University of Rennes I, IRISA  
Campus de Beaulieu,  
35042 Rennes Cedex, France  
Email: {hdrid, bcousin}@irisa.fr

Miklós Molnár  
INSA of Rennes, IRISA  
35043 Rennes Cedex, France  
Email: molnar@irisa.fr

Nasir Ghani  
ECE Department, University of New Mexico  
Albuquerque, NM 87111, USA  
Email: nghani@ece.unm.edu

**Abstract**—As optical networking deployments increase, multi-domain provisioning and survivability are becoming major concerns. A key challenge in multi-domain survivability is the scalability problem. To address this concern, various solutions have already been proposed based upon topology aggregation schemes. However, these mechanisms do not scale well with increasing domain counts and further investigation is required to develop more scalable alternatives. Along these lines, in this paper we propose to use graph partitioning techniques to solve the scalability problem in multi-domain optical networks. To demonstrate the efficiency of our method, we also extend the p-cycle concept to multi-domain settings. Overall simulation results show the efficiency of our proposed solution in terms of resource utilization and the number of p-cycle structures.

## I. INTRODUCTION

With recent advances in wavelength division multiplexing (WDM) technologies, optical networks are now capable of providing enormous bandwidth in a flexible manner where and when needed. Currently many long-haul optical fibres can readily support 140 wavelengths, yielding up to 14 Tbit/s speeds over unamplified distances of up to 160 kilometres [1]. As these infrastructures continue to scale, however, network survivability concerns are becoming more extreme owing to a wide range of potential failure scenarios, e.g., fibre cuts, node failures, software malfunctions, natural disasters, and human error. Indeed, a single failure can truly affect hundreds of lightpaths and leading to unprecedented disruption and lost revenue for network operators.

Now survivability in optical networks is a major focus area and has been well-studied for many years. The term refers to the ability of a network to continue to provide services even in the presence of a failure. To date, a variety of schemes have been studied for optical network survivability, including pre-fault protection and post-fault restoration. In general, however, network survivability becomes more and more challenging as the size of the networks increases. For example, in multi-domain settings, the sheer increase in the number of nodes coupled with the inclusion of optical-layer constraints can be very taxing, e.g., such as finite wavelength sets, wavelength continuity, and lightpath distance <sup>1</sup>.

<sup>1</sup>The maximum distance that optical signal can propagate without any bandwidth degradation

Now survivability in multi-domain optical networks entails the ability to recover end-to-end lightpaths crossing multiple domains. In general, this consists of two main tasks. The first task is to collect and maintain up-to-date network state information (e.g., link resource levels, link diversity, etc). Meanwhile, the second task is to find and reserve working and backup resources for the connections based upon the above-collected information. Now the first task is generally implemented by running routing protocols to broadcast related link-state information between network nodes, thereby allowing each to have a complete view of the network (i.e., distributed settings). However, as the network size grows, it becomes very unrealistic to broadcast complete topology information to every network node. This is further complicated when networks are delineated into multiple domains, each with its own administrative constraints and privacy issues.

Hence to deal with the scalability problem of multi-domain survivability, several solutions have been proposed. For example, various researchers have outlined topology aggregation schemes [2], [3], [4] to help represent domain-level internal topology information in a more compact manner. In turn, this reduces the amount of routing information exchanged between border nodes<sup>2</sup>. However, in the case of very large-scale networks with sizeable domain counts (e.g., such as the Internet), this method works poorly. Namely, even if every domain is aggregated, the resulting reduced topology is still very large. Take an example of a multi-domain network consisting of thousands of domains. Here, even if each domain is aggregated into a single virtual node, the state dissemination overheads are still very high. Moreover, it is noted that single virtual node aggregation offers the greatest possible reduction of advertised state information, i.e.,  $O(1)$ . As a result, aggregation is not deemed as sufficient to assure the scalability of the problem when the number of domains is large.

In light of the above, in this paper we propose to use graph partitioning techniques for improving multi-domain optical network survivability. To understand how our method works and overcomes the problems described above, we will apply it towards the protection of multi-domain optical networks using p-cycles [5]. The overall paper is organized as follows.

<sup>2</sup>Nodes that have a link to a node outside of the domain

Section 2 presents a survey of the existing work in multi-domain optical network survivability. Next, Section 3 describes the challenges associated with graph partitioning and develops a novel solution to apply this technique in multi-domain networks. The objective here is to reduce the complexity of network state exchange. Finally, the performance of the proposed scheme is studied in Section 4 for the case of p-cycle protection and conclusions and directions for future work are presented in Section 5.

## II. SURVIVABILITY IN WDM NETWORKS

Fault recovery at the lower WDM layer can yield several advantages over higher-layer (e.g., IP, Ethernet) recovery techniques, e.g., such as faster recovery, reduced overheads and complexities, etc [6]. Now there are essentially two types of recovery strategies for survivability in WDM optical networks, i.e., restoration and protection. In protection, backup resources (spare capacity) are reserved at the in conjunction with working resources. Conversely, in restoration, backup paths are searched and established dynamically after a failure on the working path. However, restoration schemes can give much longer recovery times and protection strategies. Moreover, protection schemes can also guarantee full recovery (from single or dual link failures) whereas restoration cannot, in general, if resources are not available upon a failure event. For these reasons, protection strategies are considered more germane here, particularly for survivability across multiple optical domains carrying large amounts of traffic wavelengths.

Now given that single-link failures are the most prevalent types in optical networks, many associated protection mechanisms have been proposed in the literature, see [6]. More importantly, these failure types also enable resource sharing between backup connections, thereby yielding much lower protection overheads for carriers. Leveraging this background, herein we focus on building protection solutions based upon the well-established p-cycles concept. Namely, p-cycles are chosen here as they are an effective means of deploying high-grade protection over flexible network routes. Nevertheless, our solution is generic and based upon a network partitioning strategy (detailed later in Section 3B). As such, it can readily be combined with any other protection methodology if so desired.

The concept of p-cycle protection was first introduced in [5] and has two important properties. First, a p-cycle is a fully pre-configured setup with pre-planned spare capacity. Hence when a link on the cycle fails, only the two end-point nodes of the failed link need to perform recovery actions, i.e., no switching actions are required at any intermediate nodes. Indeed, this propriety greatly improves the p-cycle recovery time and will be important in more expansive multi-domain settings. The second propriety is that the p-cycle can provide protection not only for on-cycle links (ring protection) but also for straddling links without requiring any additional spare capacity. Namely, a straddling link is defined as a link which does not belong to the p-cycle but whose both end-nodes are on the p-cycle

[5]. This property effectively reduces the required protection capacities/overheads of this technique.

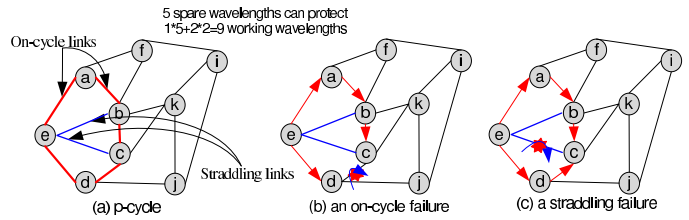


Fig. 1. Protection using p-cycles

Figure 1 presents a sample illustration of the p-cycle protection concept. Namely, in Figure 1a the node sequence a-b-c-d-e-a is termed as a p-cycle with one unit of spare wavelength capacity on each on-cycle link. Hence when the on-cycle link d-c fails, as shown in Figure 1b, the p-cycle provides one protection path (d-e-a-b-c). Meanwhile, Figure 1c also shows the case of a straddling link failure, i.e., link e-c. Here, the p-cycle can again protect two working wavelengths on this link by providing two alternate paths, i.e., e-d-c and e-a-b-c, respectively.

## III. MULTI-DOMAIN P-CYCLE SURVIVABILITY

We now present the key motivations and objectives of our proposed survivability scheme for multi-domain optical networks. Namely, the limitations of topology aggregation schemes are first analyzed. Subsequently, the key guidelines of our proposed solution are introduced in order to overcome the above limitations.

### A. Motivations

A multi-domain network is basically an amalgamation of multiple routing domains, inter-connected by inter-domain links. Here, each domain can be regarded as an independent network under its own administrative authority that has its own local rules of operation and management to provide services [8]. Now due to obvious scalability constraints and privacy limitations, complete internal topological details of a domain will generally not be available outside. As a result, no single provisioning entity can have access to complete network state information, e.g., such as knowledge of all the detailed states of wavelength usage on each link of a multi-domain optical network.

As mentioned earlier in the introduction, to address these scalability challenges, several multi-domain protection solutions have been studied, i.e., [7], [8], [9], [10], [12]. However, these primarily rely upon topology aggregation techniques to help reduce the amount of routing information exchanged between domains. Here, network domains are usually transformed into smaller "representative" domains using various graph models such as mesh, star, tree, etc. Although these schemes can hide the internal network topology of each domain, they still cannot scale for very large network sizes, e.g., thousands of domains. Namely, even if all domains are aggregated, the resultant topology state would still be

very large and entail significant routing overheads for proper propagation/updating.

In light of the above, various other strategies have been considered for multi-domain protection. For example [12] proposes a solution for protecting inter-domain links using p-cycles. Namely, this solution computes the set of p-cycles over a single-node virtual topology in which each domain is represented by a single virtual node, i.e., as shown in Figure 2b. However, this method suffers from some serious drawbacks. Foremost, is the fact that the scheme uses the entire multi-domain topology to compute the set of p-cycles. As a result, a calculated p-cycles can traverse the entire set of domains, and this will entail excessive resource management and very lengthy recovery timescales. Optical layer impairment concerns will also rise for overly-lengthy p-cycle routes here. In addition, p-cycle computation times here will be large, as they tend to increase exponentially with node counts. Meanwhile, another major drawback here is that p-cycles protecting inter-domain links are computed using a single-virtual-node topology. Namely, this topology basically summarizes a domain with multiple nodes/links into a single virtual node. As such, this high level of aggregation hides all domain-internal topological information, and this will likely impact the overall quality of computed p-cycles, i.e., redundancy is very high as the real cost of a p-cycle on physical topology will not be known.

### B. Proposed Solution

In order to resolve above the above challenges and shortcomings, we now present the main details of our proposed multi-domain p-cycle protection solution. Foremost, to overcome problems relating to protection scalability, we proposed to divide a multi-domain topology into  $k$  independent, approximately equal "sub-multi-domain" networks. Next, we propose to separately solve the problem of survivability on each of these "sub-multi-domain" networks. Furthermore, as the physical topology of multi-domain networks is very important and the graph partitioning is an NP-complete problem [5], the multi-domain topology that we propose to use in the *partitioning* process, again, is one that represents a domain as a single-virtual-node, i.e., as shown in Figure 2b.

Now after having built the  $k$  partitions, a set of p-cycles protecting each of the "sub-multi-domain" networks are computed in an independent manner. Now clearly the size of calculated p-cycles here will be limited to the number of domains existing in each partition. In addition, the amount of routing information exchanged between nodes will be reduced because associated link state information only has to be exchanged between nodes within the same partition. Moreover, associated computation times will also be drastically reduced as the independent domain-level p-cycles can be computed in a parallel manner.

Carefully note that when performing network partitioning, the inter-domain links connecting two separate partitions (i.e., links of the cut) will not be protected. To overcome this problem, we propose to further protect these links with the

links of the partition that gives the best redundancy after computing the set of p-cycles. In other words, the links of the cut to the first partition are joined and the redundancy of p-cycles which protect this new set of links is calculated. Specifically, these include the links of the partition, the links of the cut, and also the links connecting the nodes of the cut links in the other partition. The same process is then re-applied to the second partition. Now in order to illustrate this more clearly, consider the upper partition in Figure 3b. Here, the links a-f, b-f, c-g and d-h are the links of the cut, and the links a-b and c-d connect the nodes of the cut in the other partition, i.e., nodes a, b, c, d. Overall, the lower value of redundancy will allow us to decide which partition the links of the cut are associated.

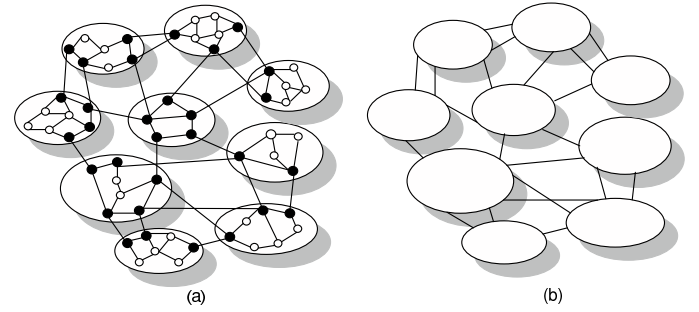


Fig. 2. Multi-domain networks

Finally, our proposed scheme also takes into account some internal topology information of each domain during p-cycle computation. This essentially addresses the earlier-cited problems related to imprecise resource state in p-cycle protection for multiple domains, i.e., in [12]. Now clearly, it is not feasible or practical to take into account the exact internal topology details of each domain here, i.e., owing to privacy concerns, computational overheads, etc. Hence we propose an intermediate solution that leverages existing topology aggregation models to condense associated domain-level state, as shown in Figure 3b. As a result, this yields a reduced "sub-multi-domain" topology on which the protection cycles can be computed [11]. Further details of the proposed graph partitioning process are now presented.

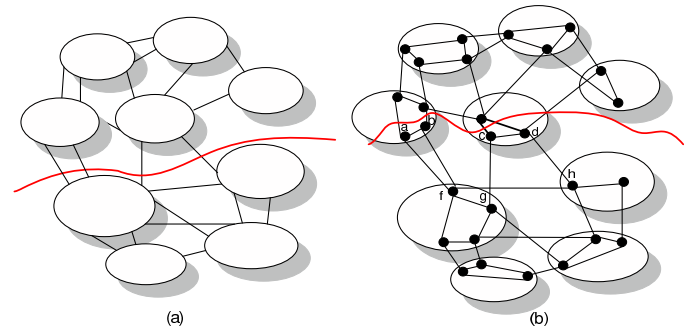


Fig. 3. Multi-domain topology partitioning

### C. Graph partitioning

Graph partitioning is a fundamental optimization problem and been widely studied both in theory and in practice. Overall, this technique is very germane and powerful means of dissecting a larger problem into a set of semi-independent sub-problems that are easier to solve. Now the main objective of graph partitioning is to divide a given graph,  $G = (\mathbf{V}, \mathbf{E})$ , into  $k$  separate parts  $(P_1, P_2, \dots, P_k)$ , i.e., where  $\mathbf{V}$  is the set of vertices and  $\mathbf{E}$  is the set of links. Generally we try to select a partition where all parts have approximately the same number of nodes (i.e.,  $|\mathbf{V}|/k$ , where  $|\mathbf{V}|$  is the number of vertices) and where the number of edges connecting different parts is minimized. Hence in our case, we divide the topology of the multi-domain network into  $k$  parts (to be detailed shortly using spectral clustering techniques [17]), with each part contains approximately  $n_d/k$  domains, where  $n_d$  is the number of domains in the multi-domain network. For example, Figure 3a shows a partition for  $k = 2$ . Overall, the  $k$  graph partitioning problem can be described as follows:

$$V = P_1 \cup P_2 \cup \dots \cup P_k \quad (1)$$

Where,

$$P_1 \cap P_2 \cap \dots \cap P_k = \emptyset \quad (2)$$

Here, the subsets are assumed to be collectively exhaustive and also mutually exclusive. Hence the total cost of the partition is defined as:

$$C = \sum_{i,j \in V, P(i) \neq P(j)} c(i,j) \quad (3)$$

Where  $c(i,j)$  is the cost of the link between node  $i$  and node  $j$  in the graph  $G$ . The objective of the above  $k$  graph partitioning problem is to minimize the cost  $C$  of the edges connecting the different parts

Overall, graph partitioning is an NP-complete problem [5], and over the years a large number of heuristic solutions have been proposed to solve this problem efficiently, see [13],[14],[16],[17]. Now for the purposes of multi-domain network partitioning herein, any of these heuristics can be applied. However, without lack of generality, we choose to apply the spectral clustering method to divide a graph into disjoint sub-graphs, as this method has many advantages. Foremost, this algorithm has been shown to yield very satisfactory results. In addition, this technique is also very simple to implement and can be solved efficiently by standard linear algebra methods. In particular, the spectral clustering method computes an eigen-vector corresponding to the smallest nonzero eigen-value of the Laplacian matrix  $L$ , where  $L = D - A$ , and  $A$  is the adjacency matrix and  $D$  is the degree matrix<sup>3</sup>.

Given a graph  $G = (V, E)$  with  $\|V\| = n$ , the degree matrix  $D$  for  $G$  is a  $n \times n$  diagonal matrix defined as :

<sup>3</sup>Degree matrix is a diagonal matrix which contains the degree of each vertex.

$$d_{i,j} = \begin{cases} \text{deg}(v_i) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

The graph is then divided into two parts according to the sign of the eigen-vector components. As a result, the graph is divided into two equal sized sub-graphs. In turn, these sub-graphs can be recursively divided to achieve higher partitioning.

### IV. SIMULATION RESULTS

The performance of the proposed multi-domain survivability solution is evaluated using simulation analysis. Specifically, a customized simulator is developed in C/C++ and run on a Dell Quadri Dual-Core Xeon processor with 4 Gb of RAM. Meanwhile, comparative integer linear programming (ILP) analysis from [5] is also done for some cases using the MATLAB toolkit. Furthermore, all tests are carried out using a commonly-used dense network topology [18], as shown in Figure 3. Namely, this topology contains 13 different domains (each region is considered as a domain) and also has bi-directional fiber links connecting any two nodes residing in different domains. All network nodes are assumed to have full wavelength conversion capabilities and each link has 32 wavelengths. Finally, traffic demands are uniformly distributed among all source-destination pairs, and it is assumed that each request is for one wavelength capacity. The working capacities on the network links are also obtained by routing each demand over the shortest path.

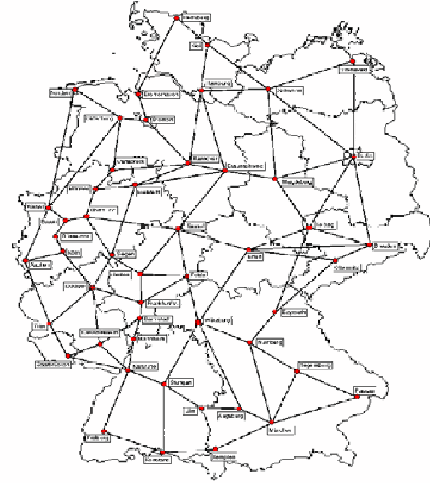


Fig. 4. Sample test network

Now in order to properly evaluate the proposed solution, several criteria are defined:

**Redundancy:** Redundancy represents a key evaluation criterion for WDM networks as it provides insight on the quality and the relative cost of network protection. Specifically, this value is defined as the ratio of spare (protection) capacity to working capacity. Hence low redundancy is more efficient than

high redundancy due to the fact that the latter requires larger amounts of spare capacity to protect against failures [15].

**Computational complexity:** Computational complexity is an important factor when evaluating any algorithm. Here it is simply defined as the time it takes for an algorithm to compute a solution. Hence it is most desirable if an algorithm can achieve a near-optimal solution within acceptable computational complexity.

**Number of p-cycle structures:** The number of p-cycle structures is another evaluation criterion here. Namely, solutions yielding a smaller number of structures are preferred as they will simplify overall network management tasks [15]. For example, consider the case where the network designers must manually configure the computed protection designs. Here, solutions with fewer structures will obviously reduce tedious configuration and maintenance tasks and reduce mis-configuration errors.

Performance analysis results with graph-partitioning are now presented for various scenarios. In the first scenario, the complete multi-domain network topology is considered when calculating the set of p-cycles. Namely, protection cycles are computed using the approach in [5], which uses an ILP model to find the optimal set of p-cycles in terms of redundancy. This scenario is used to provide a baseline case for comparing the performance of the proposed graph partitioning strategies. Meanwhile, in the second scenario the partitioning scheme of Section 3 is used to divide the multi-domain network topology into two parts. The optimal set of p-cycles are then computed for each of these sub-topologies using the same algorithm used in the previous scenario, i.e., from [5]. Finally, for the third, the above process is repeated by further dividing the multi-domain topology into four parts.

TABLE I  
REDUNDANCY

# of partitions	No partitioning	2 partitions	4 partitions
Partition Redundancy	-	89.17%	93.84%
			91.24%
		88.48%	91.67%
			91.45%
Average Redundancy	-	88.82%	92.05%

The overall redundancy results for the above four scenarios are first presented in Table 1. Meanwhile, Table 2 also summarizes the respective running times for these different solutions. Foremost, the results clearly indicate that network redundancy increases with the number of partitions chosen. Although the relative percentage is about 4% (between 2 and 4 partitions) this value effectively represents a differential in network overbuild costs and hence can translate into a large monetary "cap-ex" savings. These gains can be explained by the fact that larger partition sizes result in longer, augmented p-cycles. In turn, this lowers the spare capacity required for the set of p-cycles protecting the network and hence boosts overall redundancy. In other words, longer p-cycles tend to include more straddling links and therefore result in higher levels of protection resource efficiency.

Meanwhile, the run time results in Table 2 clearly indicate that larger partition sizes are very problematic. For example, the first scenario (single partition) does not even return a result due to the large number of nodes and working wavelengths. Conversely, the graph partitioning strategies quickly curtail run times by distributing the p-cycle computation overheads over semi-independent sub-problems that are easier to solve. For example, results show that increasing the number of partitions from 2 to 4 results in almost 2 orders magnitude lower run-times, a very notable finding.

TABLE II  
RUNNING TIME

# of partitions	No partitioning	2 partitions	4 partitions
Max running time	864000 (s)	96703 (s)	20.9 (s)
Average running time	864000 (s)	48401 (s)	8.4 (s)

Finally, the number of distinct p-cycles generated by a design is also evaluated as has a direct impact upon network management costs/complexities. These findings are shown in Table 3 and indicate that the number of distinct structures actually decrease with the number of partitions chosen in the design, e.g., by almost 50% from 2 to 4 partitions. Consequently, network partitioning can be a very attractive option for network operators, allowing them to streamline and facilitate overall network survivability provisioning.

TABLE III  
NUMBER OF P-CYCLES

# of partitions	No partitioning	2 partitions	4 partitions
# max of p-cycles	-	10	6
Average # of p-cycles	-	8,5	4

## V. CONCLUSION

This paper addresses the crucial topic of survivability in multi-domain optical networks. Owing to the sheer scale of these infrastructures, protection scalability is indeed a key challenge here. Hence in order to address this concern, we introduce a novel graph partitioning approach based upon spectral clustering techniques to help reduce the overall multi-domain network protection problem in to series of "sub-multi-domain" problems. p-cycle protection strategies are then applied to these individual entities. Detailed simulation findings show that the proposed graph partitioning approach can effectively resolve protection scalability concerns in multi-domain settings. Specifically, the scheme also achieves a very good trade-off between two differing goals, namely efficient use of backup resources and reduced algorithmic/management complexity. These results are very promising and future efforts will look at extending the work to multi-layer/multi-domain protection, e.g., IP-DWDM networks.

## REFERENCES

- [1] M. Yagi, S. Satomi and S. Ryu, *All-Optical 2R-Regenerative Interconnection Node for DPSK Polarization-Division Multiplexed Systems*, Optical Fiber Communication and the National Fiber Optic Engineers Conference, 2007.

- [2] K. Liu, K. Nahrstedt, S. Chen, Routing with Topology Abstraction in Delay-Bandwidth Sensitive Networks, IEEE/ACM Transactions on Networking, Vol. 12, No. 1, February 2004, pp. 17-29.
- [3] T. Kormaz, M. Krunz, Source-Oriented Topology Aggregation with Multiple QoS Parameters in Hierarchical Networks," ACM TOMACS, Vol. 10, No. 4, October 2000, pp. 295-325.
- [4] K. R. Bhutani, A. Battou and B. Khan, *Two Approaches for Aggregation of Peer Group Topology in Hierarchical PNNI Networks*, Int. J. Intell. Autom. Soft Comput, 2000.
- [5] W. D. Grover and D. Stamatelakis, *Cycle Oriented Distributed Preconfiguration: Ring-like Speed with Mesh-like Capacity for Self-planning Network Restoration*, International Conference on Communications, 537-543, 1998.
- [6] C. V. Saradhi and C. S.R Murthy, *Dynamic establishment of differentiated survivable lightpaths in WDM mesh network*, Computer Communications, 273-294, 2004.
- [7] D. Truon and B. Thiongane, *Dynamic routing for shared path protection in multidomain optical mesh networks*. Journal of Optical Networking, 58-74, 2006.
- [8] L. Guo, *LSSP: A novel local segment-shared protection for multi-domain optical mesh networks*, Computer Communications, 2007, pp. 1794-1801.
- [9] X. Zhang, D. Liao, S. Wang and H. Yu, *On segment shared protection for dynamic connections in multi-domain optical mesh networks*, International Journal of Electronics and Communications, 1-6, 2009.
- [10] X. Xie, W. Sun, W. Hu and J. Wang, *A shared sub-path protection strategy in multi-domain optical networks*, in Proc. Optical Fiber Communication and Optoelectronics Conference, 2007.
- [11] H. Drid, S. Lahoud, B. Cousin and M. Molnr, *A Topology Aggregation Model for Survivability in Multi-Domain Optical Networks Using p-Cycles*, 6th IFIP International Conference on Network and Parallel Computing, NPC 2009.
- [12] J. Szigeti, R. Romeral, T. Cinkler and D. Larrabeiti, *p-Cycle Protection in Multi-Domain Optical Networks*, Springer Journal of Photonic Network Communications, 2009.
- [13] G. Karypis, R. Aggarwal, V. Kumar and S. Shekhar, *Multilevel Hypergraph Partitioning: Application in VLSI Domain*. In Design Automation Conference, . IEEE/ACM, 1997.
- [14] G. Karypis and V. Kumar, *Multilevel k-way Hypergraph Partitioning*. In Design Automation Conference, pages 343-348, 1999.
- [15] H. Drid, B. Cousin, S. Lahoud and M. Molnr, *Multi-criteria p-cycle network design*, IEEE Conference on Local Computer Networks, 2008.
- [16] B.W. Kernighan and S. Lin, *An Efficient Heuristic Procedure for Partitioning Graphs*. Bell System Technical Journal, 291-307, February 1970.
- [17] U.V Luxburg, *A tutorial on spectral clustering*, Statistics and Computing, Springer US, vol. 17, no. 4, 395-416, 2007.
- [18] T-Systems International GmbH. <http://www.t-systems.de>.