# Internet Shopping Optimization Project (IShOP)

Jedrzej Musial[1], Johnatan Pecero[2], Bernabe Dorronsoro[2] and Jacek Blazewicz[1]

[1]Institute of Computing Science, Poznan University of Technology,
ul. Piotrowo 2, 60-965 Poznan, Poland
[2] Computer Science and Communications Research Unit, University of Luxembourg,
6 rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg
Jedrzej.Musial@cs.put.poznan.pl

**Abstract.** E-commerce (or even e-business) is becoming a part of modern society. The continuously growing implementation of technology (e.g., cloud computing, mobile devices - smartphones, tablets) into our daily business and administrative operations make it necessary to adapt to this inevitable evolution. One of its essential parts is Internet shopping, which becomes more and more popular with every upcoming year. Access to the service industry is now also offered through internet portals, ranging from cloud computing to translation services. Shipping costs, quantity discounts and early booking, among others, allowed for the creation of value added-services based on brokering. This project proposes innovative and realistic models for different typical online shopping operations, supported by strong mathematical and operational research fundamentals, and well balanced with lightweight computational algorithms. These models are designed in order to allow the optimization of such transactions. Finding accurate solutions to the defined problems implies both lowering customer expenses and favouring market competitiveness. Therefore, the outcome of the project will be extremely beneficial for the society; particularly taking into account that online shopping already comprises a large percentage of the actual commerce (in 2013, 50% of European consumers will be making purchases online). One of the main aims of this project is to model and formulate new advanced and realistic flavours of the Internet Shopping Optimization Problem (ISOP), considering discounts and additional conditions like price sensitive shipping costs, incomplete offers from shops, or the minimization of the total realization time, price, and delivery time functions, among others. The models will be mathematically and theoretically well founded. Moreover, the challenge of defining and addressing a multi-criteria version of the problem will be addressed too. Other important contributions will be the mapping of ISOP to other new challenges. One of them is the design of a novel business model for cloud brokering that will benefit both cloud providers and consumers. Providers will be able to easily offer their large number of services, and to get a fast answer from the market to offers (e.g., when infrastructure is under-utilized). Additionally, customers will easily benefit from offers and find the most appropriate deals for his/her needs (according to service level agreements, pricing, performance, etc.). Modelling some of these aspects and coupling it with an optimization tool for the brokering of cloud services among various providers would be a key contribution to the field. Finally, a wide set of optimization algorithms will be designed and developed for the addressed problems. They include from fast lightweight specialized heuristics to highly accurate parallel and multi-objective population-based metaheuristics. They all will

be embedded in a software framework for their practical applications, and validation.

# 1 Accordance with the Project Call - PolLux

This project deals with the e-services sector (both public and private), addressing the important market of online shopping, by modeling and optimizing very common related problems. The studies of the project are supported by a strong mathematical background, both in the definition and study of the problems, as well as in the demonstration of their NP-hardness. Additionally, efficient resolutions of such complex problems require the development of different novel IT solutions related to the effective exploration of huge search spaces. A number of both fast and highly accurate smart methods to find near-optimal solutions to the problems will be investigated. The considered topic completely matches the business service design thematic research priority of the current POLLUX call. The solution we will investigate for the internet shopping optimization problem will be a business e-service of great benefit for online customers. They will enhance competitiveness among online shops. Additionally, we propose a novel business model for cloud brokering. The cloud broker will provide his customers with the best option to deploy their services or execute their algorithms in the cloud. The choice will be made considering not only cost, but also quality of service, reliability, security, etc. Moreover, the broker will negotiate with the cloud providers unusually low prices thanks to the large market volume he generates. Cloud providers might also be highly interested in dealing with the broker special offers to get customers in some cases (e.g., when resources are being under-utilized). The project aims to design a novel business model for cloud brokering, which supports the Innovation in Services Research Priority. The engineered solutions are valid for both public and private e-services.

# 2 Description of the Project

## 2.1 Current State of the Art Including Your Relevant Previous Work

**Introduction: Electronic Commerce**
One can say that electronic commerce (e-commerce) is one of the fastest developing fields of computing science (based on operational research, combinatorial optimization). However, it is worth noticing that e-commerce is a giant hybrid built in additional areas such as logistics, economy and social sciences. E-commerce is an industry, which focuses on selling and buying products and services through web pages [1, 2]. Online shopping, fitting into a business-to-consumer (B2C) sub-category, is one of the key business activities offered over the Internet. It has become increasingly popular over the past decade. U.S. e-commerce market sales amounted to 289 billion USD in 2012, up from 256 billion USD in 2011. In Poland, Internet retailing in goods amounted to 2 billion Euros in 2010. Luxembourg is a strategic place for e-commerce and online shopping with the key players like Amazon, Skype, iTunes, PayPal, eBay, and many other leading ICT companies [3]. This attractiveness is mainly due to the state-of-the-art infrastructure, telecommunication network (e.g. TeraLink network), one of the best

data center parks in EU to support cloud computing, trust culture (with LuxTrust Luxembourg being the first European country to provide online certification for the private sector) and one of the lowest VATs in the EU.

**Internet Shopping Optimization**

Price comparison sites (e.g., Google Shopping, Shopping.com, PriceGrabber, or Kelkoo are among the most popular shopping engines in EU) are search tools designed to give price information from many retailers through a single portal. However, current price ranking solutions target only single product buying. A significant percentage of price comparison websites perform sub-optimally in one of their major functions: presenting prices and considering multiple products.

Motivated by the problem of buying multiple products from different e-commerce web sites, we proposed a combinatorial optimization problem where we have a set of products $N = 1, \ldots, n$ and a consumer who wishes to buy all products in $N$ from a set of Internet stores $M = 1, \ldots, m$ at the minimum final price, subject to the price, availability, and discount constraints at each store [4–6] (For a double discounting function one can follow [7]). For $i = 1, \ldots, m$, let $S_i$ denotes the decision variable that corresponds to the subset of products selected by the consumer from store $i$. The corresponding optimization problem, which the authors refers to as the ISOP (Internet Shopping Optimization Problem), is given by:

$$\min \quad \sum_{i=1}^{m} (d_i y_i + \sum_{j \in N_i} p_{ij} x_{ij}),$$

$$\text{s.t.} \sum_{i \in M_j} x_{ij} = 1, \; j = 1, \ldots, n,$$

$$0 \leq x_{ij} \leq y_i, \; i = 1, \ldots, m, \; j = 1, \ldots, n,$$

$$x_{ij} \in \{0, 1\}, \; y_i \in \{0, 1\}, \; i = 1, \ldots, m, \; j = 1, \ldots, n.$$

where $d_i$ denotes the delivery price from store $i$ and $N_i \subseteq N$ denotes the set of products available at store $i$. Moreover, for each $i$ and $j$, $p_{ij}$ denotes the price of product $j$ in store $i$. The function $f_i : R^+ \to R^+$ denotes the discount function associated with store i (which is a function of the delivery and total price of the products).

During previous research different versions (specializations) of ISOP were examined. All of them are described in journal publications as well as presented in a number of international conferences [8, 9]. The main stress was placed on the most complicated version of the problem (the most realistic at the same time) where we deal with the discounting functions for every store.

It is worth noticing that there are some similarities with the well-known Facility Location Problem (FLP) [10]. The main characteristics of the FLP are space, the metric, given customer locations and given or not given positions for facility locations. A traditional FLP is to open a number of facilities in arbitrary positions of the space (continuous problem) or in a subset of given positions (discrete problem) and to assign customers to the opened facilities so that the sum of opening costs and costs related to

the distances between customer locations and their corresponding facility locations is minimized.

Discussions of FLPs can be found in [11, 12, 10, 13, 14]. The traditional discrete FLP is NP-hard [15, 16] in the strong sense. It was also proved that ISOP without any discounts and flat shipping rates is NP-hard as well [4]. Note, however, that the general problem Basic ISOP with price sensitive discounts cannot be treated as a traditional discrete FLP because there is no evident motivation for a discount on the cumulative cost in the sense of distances. It can be noticed that this problem and problem Basic ISOP are not sub-cases of one another, while the traditional discrete FLP is a special case of any of these problems.

Internet shopping problems were under interest of the working group already for some time. As a result of this work one can mention several journal publications [4–6]. Moreover, Internet shopping problems, mathematical formulations, complexity discussions, as well as algorithms propositions and some first computational tests were done within a Ph.D. thesis by an applicant entitled "Applications of Combinatorial Optimization for Online Shopping" [17]. What is surely worth noticing is that the thesis was written in a co-tutelle programme between Poznan University of Technology and the University of Luxembourg. Furthermore, good collaboration between both universities groups brought about another publication [5, 6]. Results of the collaboration have already been presented during some OR / combinatorial optimization conferences (ECCO 2012, EURO 2012) and this trend will certainly be continued (already accepted for presentations this year during ECCO 2013, MISTA 2013, OR 2013). The main contributions of the previous work within Internet shopping problems are the following:

- Mathematical formulation of the Internet Shopping Optimization Problem (ISOP),
- providing the proof that the ISOP problem is NP-hard by a polynomial transformation from the Exact Cover by 3-Sets (X3C), which is a well-known NP-complete problem,
- detailed literature analysis with a special attention paid to Facility Location and Knapsack Problems,
- algorithm design  a preliminary version of the algorithm solving basic ISOP was prepared.

**Cloud and Cloud Brokering**

Cloud computing [18, 19] is, undoubtedly, one of the main existing computing paradigms nowadays. In the last years, it raised the interest of both academic and industrial worlds thanks to their interesting properties, such as elasticity, flexibility, or computational power, among many others. Cloud computing provides a stack composed of different kinds of services to users [20]: Infrastructure as a Service (IaaS), dealing with resources as servers, storage, or networks; Platform as a Service (PaaS), which provide an operating system as well as a set of tools and services to the user; or Software as a Service (SaaS), that allows providers to grant customers with access to licensed software.

Many different public and private clouds are arising in the last years [21]. They all have distinct features, making it difficult for users to find the best choice among all the existing services offered by the Cloud Service Providers (CSPs) [22]. The reason is that cloud users must define their specific requirements in terms of services and application

deployment [23]. These specific requirements can be categorized into measurable features such as cost or resources and user-experience features such as response time or performance (based on the feedback from the cloud users experience about the service they purchased from the CSPs). These parameters must be collected from the data log of the VMs/Services of the cloud user on-running systems using machine learning system to compare with promised QoS of the CSPs.

An example of the measurable features is presented in the study of Chaisiri et al. [24]. The authors pointed out that the pricing model offered by CSPs includes on-demand and reserved prices. However, cloud users also need to take into consideration the on-running cost for the instances that they purchased from CSPs. So there are three phases for the life-cycle of VM management in CC: reservation, on-running and on-demand phase [25, 26]. Based on this model, the authors presented the differences between private and public CC not only in the price plan but also in the resource limitation.

In cloud computing, CSPs must offer their services to cloud users with the promised SLA and Quality of Service (QoS). There are a few differences in cost or plan between CSPs [27]. The characteristics of the services offered by CSPs can be defined as feature and non-feature characteristics. The feature characteristics are price model (average monthly price), resource plan and SLA/QoS. The non-feature characteristics can be defined as virtualization, performance, reliability and security. CSPs can also be categorized with respect to the service model: SaaS, PaaS and IaaS, or the QoS and SLA for the services they provide.

The figure of cloud broker [28] arises as an intermediary entity between cloud providers and users to help the latter ones in the process of choosing the most appropriate services among those offered by the different CSPs, according to their particular needs. There are different services that cloud brokers can provide, from simply finding the best deals among a set of clouds for the user requirements to defining the best possible design to deploy the user's application in the cloud [28]. Additionally, the cloud broker can consider the QoS and SLA from the CSPs as parameters to compare with the results from the historical data or log analysis based on the on-running services of the cloud users, in order to ensure high QoS.

There are numerous available studies in the field of research for the cloud broker. In the research conducted by Spillner et al. [29], the authors show that the cloud broker is not only the interface to manage the virtualized resources between the cloud provider and the cloud user but also can help to bring the unused resources from the cloud users to reuse in the cloud. Therefore, they proposed to define a nested VM in which other VMs from multi-cloud providers are referred to as sub-VMs. The concept of nested VMs helps the cloud broker to deal with a variety of VMs from multi-cloud providers. In these sense, there is a business model in which the cloud broker buys reserved instances from a number of clouds and then sublet them as an on-demand basis to the users at cheaper prices than those of the cloud providers. This is profitable thanks to the high price difference between on-demand and reserved VMs [30, 31].

Usha et al. [32] also proposed a framework for cloud brokerage service. The framework schedules the cloud resources by considering both the multi-criteria objectives from cloud users and cloud providers. Their proposed model based on the QoS parame-

ters includes: response time and the throughput. The optimization problem is defined as the multi-criteria optimization so that the author proposed to use the Pareto technique (the Pareto front) to find the optimal solutions.

Carpentier et al. [33] presented the CompatibleOne, an open-source framework for the cloud brokerage service to work with a multi-cloud environment. Two important components were proposed in the CompatibleOne: (a) the energy-monitoring module and (b) the module to interact with the cloud monitoring systems.

Cloud brokers are intermediary entities between the cloud providers and their customers that provide different services to the users of the cloud. From the previous literature review we can notice that different service management problems have been considered from the cloud brokering point of view. We define in this project a service management (cloud brokering) problem by considering a customer requiring a set of (different) services offered on the cloud by a cloud provider or by a set of private and public cloud providers with an associated service cost (different for each service provider). The aim is to minimize the service cost from the customer point of view and maximize resource utilization from the service provider point of view. The service management problem can be adopted by cloud brokers to provide a better service to the customers and deal low prices with cloud providers, who at the same time might benefit from such a broker negotiating with customer demands in case of under- or over-utilization of resources. The simplified cloud brokering problem can be reduced to the ISOP problem such that, algorithms and approaches developed to solve ISOP can be applied to the cloud brokering problem. However, advanced models of cloud brokering need to consider carefully different service level agreements between customers and service providers, as well as to deal with different quality of services (services requirements, resources availability, discounts, etc.) that will affect the optimization of the problem.

### 2.2 Project Objectives and Contribution to Knowledge Development

The project is expected to contribute to the existing knowledge in different fields, as it is outlined in Figure 1.

From the application point of view, we expect outstanding contributions related to the problem of providing optimal planning for internet shopping (the ISOP problem). This is an important problem for society, especially nowadays that internet shopping is becoming more and more popular and convenient, that has not yet been adequately addressed in the scientific literature. This project will focus on the modeling of the problem, with very strong mathematical foundations, on the study of its main features, and the development of novel highly efficient heuristics to provide accurate solutions. Therefore, important contributions are expected in the field of operational research. One important objective of this project is to provide other researchers with a solid highly realistic model for this important problem. We expect many different research groups worldwide will be interested in the research line we plan to open thanks to this project. Several new subproblems of the ISOP will be defined depending on particular parameters. To be more precise, we plan to:

- Analyze requirements and describe extended models for the ISOP problem.

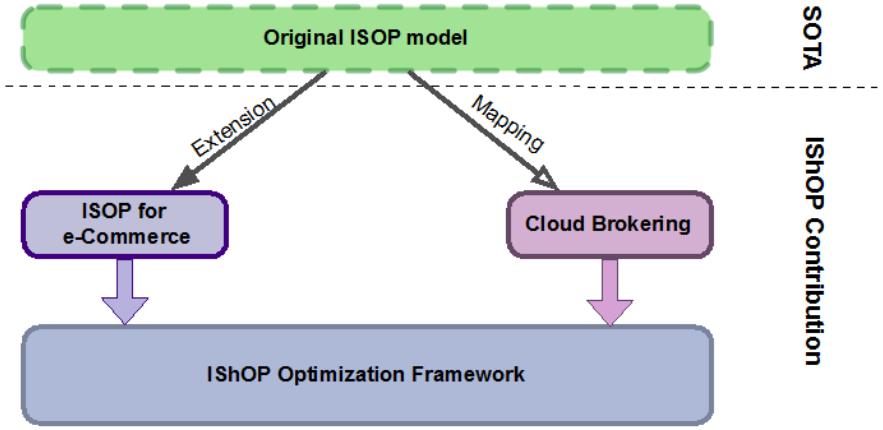


**Fig. 1.** Summary of the main contributions of IShOP project.

- Formulate Internet Shopping Optimization Problem with discounts and with additional conditions:
  - Price sensitive shipping cost  idea that follows discounting function  a few thresholds with different shipping costs,
  - Minimum delivery times,
  - Incomplete shopping lists realization.
- Provide the models with mathematical and theoretical foundation.
- Prove NP-hardness.
- Analyze decision-aid multi-criteria ISOP. ISOP will be enhanced with total realization time minimization of the total function (utility function):
  - price and delivery times functions,
  - decision-aid recommender function. Only higher recommendation rank shops influence of the recommendation rank on the total price.

For these problems new algorithms will be prepared and tested. The results will be presented in several journal publications. A software platform aimed at testing the new approaches will be also developed.

As an attempt to disseminate the modeled ISOP problem in different fields and to show its generality, we plan to map it into the brokering problem for cloud computing, in order to help the broker to manage the huge existing offer of services. We are particularly interested in finding the best offers from a number of cloud providers that suit the requirements of the customers. Moreover, the modeled cloud broker will provide additional advantages, such as negotiating lower prices with cloud providers thanks to the large market volume they generate. Cloud providers will also benefit from the broker, since they will be able to offer the broker good terns to get customers in case of underutilization of the resources, or to direct customers to other providers in peak times when the resources are overutilized. This will be done in a transparent way to the customer, so the reputation of the provider will not be damaged. The novel cloud services management agent we will work on is applicable to any cloud broker in order to make it much more advanced that the ones existing in the literature, and therefore it will be

a groundbreaking result. The scientific literature about cloud brokering is large, and it is one of the hot topics in the cloud computing field, because it is considered as one of the mechanisms that will allow the federation and interoperability of clouds in the future. Additionally, cloud brokering is a novel business model, and many new companies appeared in the market in the last years.

The proposed optimization for cloud broker will be based on the development of the ISOP, and it should deal with conflicting objectives that will benefit both cloud users and CSPs. This is therefore a multi-objective optimization problem that will be solved with simple heuristics [34], sequential and parallel local search algorithms [35, 36], and Evolutionary Algorithms [37–41]. The algorithms will look for the optimal Pareto sets of solutions that satisfy both cloud users and CSPs.

The optimization problem of the cloud users and the CSPs includes both the measurable and user-experience parameters. The measurable parameters are cost and resource placement. While the user-experience parameters contain response time, performance and security. The term user-experience means that these parameters will be based on the user experience about the services they purchased from the CSPs, these parameters will be included into a machine learning model with scoring function about the provided services of CSPs. The proposed mathematical model is studied to deal with the optimization problem that contains measurable parameters: cost and virtual machine (VM) placement. Moreover, a machine learning technique will be used to deal with user-experience parameters from the results of historical data or log based on the on-running services of cloud users.

The concept of the proposed cloud brokering methods can be defined by characteristics as:

- Interface: the cloud broker is the interface between cloud users and CSPs,
- the cloud broker plays an important role to negotiate in order to balance the optimization objectives of cloud users and CSPs. It is obvious that there are conflicts between both parties objectives. Therefore the cloud broker offers a multi-objective optimization method to find optimal tradeoff solutions from the Pareto front,
- choosing and assigning tasks/services from the requests of cloud users to deploy on the services of CSPs,
- managing VM/Service repository with service classification (the VMs or services run the same task),
- managing cloud services stack from the available CSPs.

In the design concept of the cloud broker, there is a VM/Service repository that stores VM or service specific requirements and cloud service stack stores details of available CSPs. The VM/Service repository stores parameters including:

- Measurable parameters:
  - Resource vector v = CPU, RAM, storage, network inbound/outbound data
  - Cost: pricing model reservation, on-demand, on-running
  - Service model: SaaS, PaaS, IaaS
- User-experience parameters:
  - Security
  - Plan (Small/Medium/High/etc.)

- User-location
- Time peak/low schedule

In order to successfully carry out our investigations for this project, we will need to make other important contributions in a number of different fields as decision aid, multi-criteria decision making, multi-objective optimization or parallel optimization algorithms, among others.

### 2.3 Methods and Approach

Problems to be considered in this proposal, involve many constraints and objective functions, which are defined on sets of discrete objects rather than on continuous domains. Problems of this kind have a combinatorial nature. Very often problems of this nature are hard to solve because they require verifying a number of possible solutions, which grow exponentially with the problem size. Since exponential functions grow almost explosively, it is practically impossible to solve such problems by full enumeration of possible solutions. Such approaches are called exact. Unfortunately, they are practically ineffective. On the other hand, effective algorithms are the ones in which execution time can be bounded by a polynomial in the input size. Unfortunately, according to the current state of knowledge, there exist hard combinatorial problems for which no effective, or in other words polynomial, algorithms exist.

The theory of computational complexity provides methodology of dealing with hard combinatorial problems. In particular, the methods of recognizing hard combinatorial problems, discerning their different classes and providing types of usable algorithms for certain classes. The first step in analyzing a combinatorial problem is to verify if it is solvable in polynomial time. If not, then a proof should be presented that the problem belongs to a class of hard problems such as NP-complete or NP-hard problems. In such a case the problem can be dealt with in a few ways: an exponential running time may be proposed, however with a limited applicability. A heuristic may be also proposed. Heuristics are a group of algorithms that run in polynomial time, and hence they are effective and provide correct solutions, yet without a guarantee of optimality. Heuristic algorithms can be analyzed with the goal of providing guarantees of solution quality, e.g. that the ratio of the value of the solution provided by the heuristic and the optimum objective value is bounded.

Problems considered in this project are also hard combinatorial optimization problems. Therefore, they will be analyzed according to the above methodology. First, the complexity class of the particular problem must be determined. In a best case the problem may be solvable by an effective (polynomial) algorithm. Yet, due to the nature of the solved problems and their computational hardness it is more likely that other algorithms will be applied: exact, exponential-time or heuristic, polynomial-time methods.

Efficient solving of practical problems, such as internet shopping and cloud brokering require, first of all, precise and concise formulation. The systematic research will be preceded by proposing mathematical models of considered problems, which reflect the real world environment in the set of variables and constraints. The general models allow collecting the most important components and features describing the problem and they are the basis for some relaxations and direct the future research. Depending

on the ranges of values, their distributions and the impact of particular constraints on some variants of the problems might be distinguished, varying in their complexity. It would be interesting to investigate computationally easy cases, which polynomial time algorithms, might be used as sub-procedures allowing for faster solving of more general cases. Moreover, the analysis of the mathematical models should support searching for similarities between the considered problems and closely related optimization problems, such as the travelling salesman or purchaser problems, in order to incorporate some methods developed for these cases. Since real world problems are usually computationally hard, the research will focus on constructing heuristic (efficient in quality) and exact (efficient in time) algorithms, after formal proving NP-hardness of the analyzed cases. For algorithms used, e.g., in on-line environment for internet shopping - the time complexity is a crucial factor, since it is necessary to ensure acceptable response time. For methods used e.g. for advertisement placement, the quality of solutions is more important than their run times. Depending on the considered case, proper algorithmic strategies will be used: fast simple heuristics or more sophisticated and consequently slower metaheuristics, and even branch and bound methods. In some cases, efficient dynamic programming approaches will be taken into account, which should give optimal solutions in reasonable time. Mathematical programming formulations might be also useful for incorporating the models into commercial solvers such as e.g. CPLEX, which might serve as source of reference solutions.

The above discussion shows that three groups of algorithms may be applied to solve hard combinatorial optimization problems:

1. Exact algorithms. The main feature of exact algorithms is that they search the space of possible solutions exhaustively. It means that in the worst case all possible solutions may be visited. Therefore, in the worst case, execution time of such algorithms may be prohibitively high. However, there are techniques allowing for pruning the search space. For example, in branch-and-bound algorithms the space of solutions is represented as a search tree. The tree is pruned by using bounds on values of solutions that may be derived from a given node of the tree. Branch-and-bound algorithms were applied in the past with a relative success. Hence, wisely crafted exact algorithm may be practically usable.
2. Dedicated heuristics are a group of methods, which are created to solve just one combinatorial optimization problem. Due to different characteristic of the problems it is not possible to transfer such heuristic to a different problem. Still, by their polynomial running time they are very often applied to solve combinatorial optimization problems.
3. Metaheuristics are a group of general purpose methods which can be adapted to many combinatorial optimization problems. Metaheuristics provide general schemes of conducting search in the space of combinatorial problems. Since they are very general, it is necessary to provide an interface to the actual combinatorial optimization problem. Example metaheuristics are simulated annealing, tabu search, genetic search, ant systems and many others. This type of algorithmic approach to combinatorial optimization was very successful and has been extensively studied in recent years.

To tackle the complex NP-hard problems defined in the project, advanced optimization methods need to be developed. High effort will be made to design highly efficient and accurate heuristic algorithms to solve the considered problems. A guarantee of the performance of such tools will be provided, for instance by validating the results they provide with the optimal solution, computed by exact methods, when possible. The possibility of designing algorithms that guarantee a given maximum error on the solution from the optimum will be explored too. Our previous experience on the design of optimization algorithms for problems as tasks scheduling in high performance centers [39, 38, 42, 43, 23, 44, 22] or our preliminary ISOP problem [4–6, 8, 9, 17] will constitute an asset.

The considered problems that we study in this project are multi-objective in nature. This means that they usually require more than one objective to optimize, none more important than the other, and that are usually in conflict (for a given optimal solution, increasing the quality of one of the objectives leads to worsening some of the others). We therefore plan to apply advanced novel multi-objective algorithms to find a wide and diverse range of accurate tradeoff solutions (i.e., a Pareto front approximation) in order to study the shape of such set of solutions. This study is very useful because it will allow us to get important knowledge on the problem such as knowing which kind of tradeoff solutions we are more interested in (i.e. which region of the Pareto front approximation), and then we can design efficient heuristics to search that area. Additionally, parallel algorithms (both for multi-objective and single-objective optimization) will be considered to speed-up the optimization process in the case they are required.

Moreover, the important part of the project (Software Framework) will be to implement and experimentally validate a theory-driven online model in realistic testbeds to demonstrate the performance improvement over currently used techniques. The task is challenging and requires concentrate a large effort work with the algorithms design, computational experiment, and data collecting engine preparation. The latest mentioned task is a priority to perform the best possible quality of computational experiments. These will confirm the highest quality of proposed algorithms. Model load should be perfectly balanced since it will be accessible online. During the mentioned package we plan to use several approaches to achieve our goals:

- Develop efficient lightweight heuristic based algorithms, considering approximation factors, greedy based algorithms.
- Investigate pseudo-parallel cellular based optimization algorithms (shorter computation time).
- Mathematical analysis of the performance of the algorithms.
- Computational tests of the developed algorithms.
- Research on smart algorithms to find highly accurate solutions and target large problem instances.

Performance of computer algorithms can be assessed in many different ways. Yet, the most important aspects are related to the resources consumed by the algorithm and to the quality of the results the algorithm provides. The resources may have different nature: memory, processors, time. The quality of algorithm output may be also measured in many ways: e.g. as the correctness of the output, frequency of errors, etc.

In the context of combinatorial optimization the key algorithm performance indicators are runtime and distance of the solution from the optimum. Runtime, or the algorithm execution time, is very important because exact algorithms, such as branch-and-bound (B&B), require exponential time in the size of the input in the worst case. However, successful B&B implementations are capable of mitigating this exponential growth. Furthermore, it is a very common situation that a trade-off between solution quality and the runtime exists, especially in the case of metaheuristics. Thus, some algorithms may be more effective in trading time for the solution quality.

Algorithm execution time and solution quality distance from the optimum may be studied in various ways. One possible approach is analytical. In this case, the order of the computational complexity of the algorithm is searched for, as an indicator of the algorithm temporal cost. The computational complexity of an algorithm is a function of the algorithm execution time in the size of the input. The quality of the solutions is assessed by analytically providing evidence (e.g. a proof) that the distance of the solution quality from the optimum, in the worst or average case, is bounded.

Performance of algorithms may also be studied experimentally. In this case, the algorithm is tested on benchmark instances. Its runtime and solution quality is compared against other algorithms. Note that certain algorithms may be better on some test cases and worse on others. Consequently, experimental analysis may be inconclusive when the criteria of the comparison are not strictly defined. Quite often statistical analysis of the results is accepted as a method resolving such problems.

## 3 Description of the Project Plan and Work Packages

IShOP project is divided into 5 Work Packages, organized as it is shown in the Figure 2.

Each of the Work Packages tackles a full topic. The tasks in the different Work Packages are designed to be iterative and at a given level most of them can be done in parallel. WP1 and WP5 are pure management tasks to focus on proper project guidance and best possible presentation of the results (workshops, conferences, publications, reports, websites, etc.).

We combine mathematical, theoretical research work/formulations/modeling in WP2 and WP3 with practical use of them (as preparation on test applications in Software Framework WP4). Algorithm design and optimization are extremely important and we will focus on them in WP4 as well as partially in the previous WPs. All WPs will provide results that will be important contributions to e-commerce and, more generally, operational research, optimization and computer science. Objectives, deliverables and milestones were defined for the different WPs, and they are described in detail in each of the Work Packages.

What is surely worth noticing is that the project will promote the collaboration between Poznan and Luxembourg into higher level of quality. Running IShOP project standalone on each university would definitely not be feasible. The key factor was to combine the best possible top-class specialists in each field and make them WP leaders. Moreover, the whole project team is carefully picked to complement their skills and to lower the risk of failure to the minimum. Furthermore, each WP group is balanced with

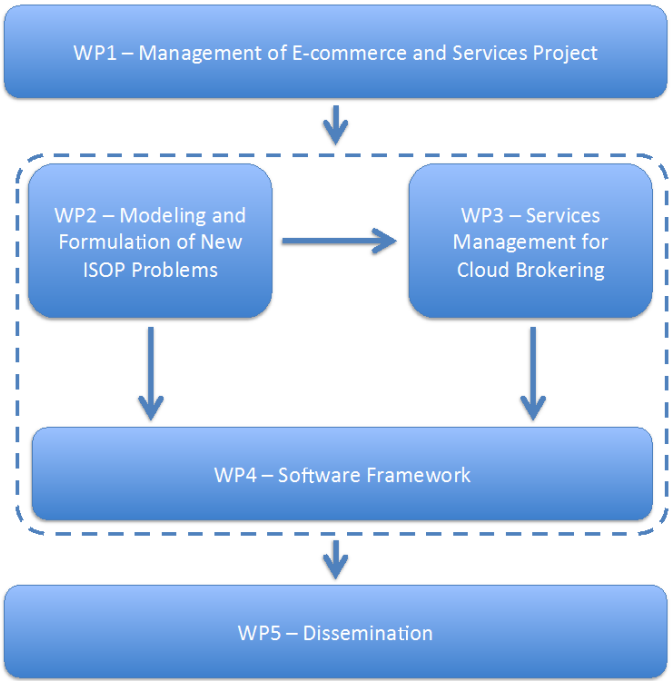


**Fig. 2.** Project plan: packet interdependencies.

both Polish and Luxembourgish scientists in order to combine their expertise for successfully accomplishing the objectives (each team represents different specific skills). That will also allow to positively stimulate one another and provide new ideas. According to our experience, this approach is much better for the project quality and risk management than working in hermetic groups. Some of the members involved in the project from the two groups (UL and PUT) are already used to work together, and some common publications proof that.

The Project PI is Prof. Dr. Jacek Blazewicz. His scientific experience (over 340 papers in outstanding journals, numerous international awards, and prestigious positions in research societies, H-index=26) is a guarantee of a high chance for a successful output of the grant. Prof. Dr. Pascal Bouvry (from UL) was the quality director of a major software development company for CLES, R&D director and general manager for SDC company, as well as PI for a number of research projects. His managerial and scientific expertise will be key for IShOP.

## 4 Project Outputs

The first contribution of this project will be a formulation of several new optimization problems. All of these problems arise from practical concerns originating from e-commerce, and they will be deeply analyzed from a practical point of view and well defined with the use of formal mathematical formulation. Therefore, we expect they should open new interesting research areas. Proofs of NP-hardness will be provided.

For multicriteria problems or problems concerning optimization of several goals, different objective functions will be proposed. Whenever possible, approximation factors will be considered. The above issues will offer a very good starting point for further research, e.g. algorithmic, but also deeply practical.

The second output of this project will be the creation of the strong research group conducting research in the area of algorithms and services for e-commerce and related topics, such as cloud service optimization problems and website optimization problems, originating from cooperating researchers from both institutions. Currently, these topics are covered almost isolated at both universities, mostly by sole researchers. Still, results achieved so far show that this area is interesting and important, only very underdeveloped. The new situation would join experienced researchers, with wide theoretical knowledge of algorithms and methods, with young ones (even students) with most up to date understanding of e-commerce current trends and ideas, and thus allow to reach the critical mass. This would also benefit in researcher training for one PhD student and seven postdocs, including some freshly ones after their dissertation, as well as highly experienced ones. Both partners would greatly benefit from this project in terms of development and visibility. A non-disclosure agreement will be signed between the two parties composing the consortium after the project is granted, in order to ensure a correct management of the intellectual property rights of the new achievements.

The foreseen publication output of this project is composed by at least three peer reviewed research papers (related to WP2, WP3 and WP4) published in good quality journals from the ISI list and over 6 papers published in top conferences (peer reviewed). We strongly believe that research results, especially publicly funded, should be as accessible as possible, thus we plan all the papers to be published in the open access, when possible. Choice of journal with good impact factor will lead to subscription model, fees for open access articles will be paid.

Both teams, PUT and UL, will have the opportunity to disseminate the knowledge gained with the project to the classrooms (highly related subjects are taught by the researchers involved in the project at PUT and LU, both at master and Ph.D. programs). There will also be opportunities for some students to directly participate in the project (student contracts are planned). Previous experience indicates that students show more interest when knowledge is illustrated with practical applications, of course, with well-founded theoretical background. Additionally, the results obtained from this project will be of high interest for the two institutions to compliment programs already on-going at both sites. The methodology developed and used in this project to analyze theoretical performance guarantees and approximation factors for optimization algorithms is also interesting for the goals of the research teams of the two institutions and valuable for other research projects in fields related to security, bioinformatics, Grid and cloud computing, in the frame of academic and industrial research projects (e.g., SUPER Node II in partnership with MixVoIP company, a satellite payload optimization project in partnership with SES satellite company, ECO-CLOUD FNR project, or INDECT FP7 project).

We expect key knowledge to be generated and developed in the framework of this research work. The project will thus contribute to the advancement of state-of-the-art knowledge production and its dissemination through peer reviewed publications, work-

shops, and journal special issues organization on subjects related to algorithms, optimization, services for e-commerce, and cloud computing. Knowledge acquired by different postdocs working on the project will be also very valuable when looking for academic positions or key industrial positions. This will also operate sibling and dissemination of knowledge.

We are convinced that by combining the experience of the groups and the convergence of interests for research, design, and implementation, this project will make major advances in the areas of operation research, optimization, cloud computing, and e-business. Among the results expected in this and future collaborations, the following can be also emphasized:

- Training of high-level researchers who can support the postgraduate computer science program in Poland and Luxembourg.
- Establishment of favourable conditions for future projects that consolidate the Poland-Luxembourg relation in the areas of Operational Research and Computer Science.
- Accumulating resources, expertise and previous scientific results of the two teams to provide solutions to todays e-commerce (internet shopping) and cloud computing (cloud brokering) problems.
- Publishing articles in highly recognized and prestigious international journals and conferences.

IShOP will achieve the major goal of the POLLUX program that is covered by fostering scientific quality in priority domains such as business service design, by optimizing business services strategies related to Internet shopping, cloud brokering, and e-commerce websites. Internet shopping, which can be considered a sub-category of e-commerce as it predominantly refers to business-to-consumer (B2C) transactions such as online retail or online auctions, is one of the key business activities offered over the Internet. At a time of economic crisis when consumers and businesses are under tight financial constraints, there is a shift towards e-commerce, and it is expected to have more and more importance in the world economy in the coming years.

The IShOP consortium strongly believes in dissemination of research findings through journal and conference publications. They will occur as a result of the training of Masters, PhD, and postdoctoral researchers. The IShOP consortium is aware of the growing need to establish workable relationships with leading researchers. The workshop organization will help to involve highly recognized scientific researchers from the community worldwide.

The consortium aims to achieve high impact and excellence well beyond the academic community by collaborating with external partners such as industry partners, government bodies, institutions and international actors. The intended result of this project is to seek cooperation with a business that could use elaborated ISOP results. At the current state of art this is not possible, as models are too simplified, and they exclude too many real life situations. The improvements are meant to meet the requirements of large online shops (like Amazon) or price comparison sites (like Google Shopping, Ciao!, Ceneo, etc.). Industrial session organization will strongly support relations with external partners, as has been demonstrated by the success of Business Meets Research forum in Luxembourg, to show the expertise, capabilities and innovative solutions of the IShOP consortium.

Cloud computing is positioning itself as a new emerging platform for delivering information infrastructures and resources as IT services on a pay-as-you-go basis to streamline business processes. It has evoked a high degree of interest internationally, with many challenges such as security, privacy and resource optimization remaining open. Luxembourg is one of the best data center parks in Europe with more than 15 data centers and more than 35,000 m2 for IT rooms. The country is significantly increasing its level of cloud computing expertise and extending this skill Europe-wide. This knowledge acts as a springboard for the future worldwide developments of the Luxembourg computing cloud. However, due to the vast diversity in the available cloud services, it has become difficult for the customers to decide whose services they should use and what is the basis for their selection. The IShOP consortium will contribute to alleviate this problem by developing models, algorithms and mechanisms to the cloud brokering problem. Such models and algorithms can make a significant impact and will create healthy competition among cloud providers to satisfy their service level agreement and meet customers quality of service requirements.

IShOP will serve as the basis for educational material at Master level (module in the problem solving and optimization lectures at UL as well as algorithm design and electronic commerce dedicated lectures at PUT) and PhD level (seminars for the SnT, for the CSC research unit and ILIAS laboratory and the Doctoral School at PUT). The project will also develop and improve the transfer of knowledge between Poland and Luxembourg.

## Acknowledgements

This project is supported by the FNR (Luxembourg) and NCBiR (Poland), through IShOP project, INTER/POLLUX/13/6466384.

## References

1. Hagel III, J.: Net gain: Expanding markets through virtual communities. Journal of Interactive Marketing 13 (1999) 55–65
2. Timmers, P.: Business models for electronic markets. Electronic Markets 8 (1998) 3–8
3. Coopers, P.: Your access to european market. [online] (2009) http://www.pwc.com/lu/en/invest-in-luxembourg/docs/pwc-publ-lux-where-else.pdf.
4. Blazewicz, J., Kovalyov, M.Y., Musial, J., Urbanski, A., Wojciechowski, A.: Internet shopping optimization problem. Int. J. Appl. Math. Comput. Sci. 20 (2010) 385–390
5. Blazewicz, J., Bouvry, P., Kovalyov, M.Y., Musial, J.: Internet shopping with price sensitive discounts. 4OR-Q J Oper Res 12 (2014) 35–48
6. Blazewicz, J., Bouvry, P., Kovalyov, M. Y., Musial, J.: Erratum to: Internet shopping with price-sensitive discounts. 4OR-Q J Oper Res (2014) online first.
7. Blazewicz, J., Cheriere, N., Dutot, P. F., Musial, J., Trystram, D.: Novel dual discounting functions for the Internet shopping optimization problem: new algorithms. J Sched (2014) online first.
8. Wojciechowski, A., Musial, J.: Towards optimal multi-item shopping basket management: Heuristic approach. In Meersman, R., Dillon, T., Herrero, P., eds.: On the Move to Meaningful Internet Systems: OTM 2010 Workshops. Volume 6428 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2010) 349–357

9. Blazewicz, J., Musial, J.: E-Commerce Evaluation – Multi-Item Internet Shopping. Optimization and Heuristic Algorithms. In Hu, B., Morasch, K., Pickl, S., Siegle, M., eds.: Operations Research Proceedings 2010. Operations Research Proceedings. Springer Berlin Heidelberg (2011) 149–154
10. Revelle, C., Eiselt, H., Daskin, M.: A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research 184 (2008) 817–848
11. Krarup, J., Pisinger, D., Plastriab, F.: Discrete location problems with push-pull objectives. Discrete Applied Mathematics 123 (2002) 363–378
12. Eiselt, H., Sandblom, C. L.: Decision analysis, location models, and scheduling problems. Springer (2004)
13. Melo, M., Nickel, S., Saldanha-da Gama, F.: Facility location and supply chain management - a review. European Journal of Operational Research 196 (2009) 401–412
14. Iyigun, C., Ben-Israel, A.: A generalized weiszfeld method for the multi-facility location problem. Oper. Res. Lett. 38 (2010) 207–214
15. Garey, M., Johnson, D.: Computers and Intractability: A Guide to the Theory of NP-Completeness. New York, Freeman (1979)
16. Blazewicz, J.: Złożoność obliczeniowa problemów kombinatorycznych. Warszawa: Wydawnictwa Naukowo-Techniczne (1988)
17. Musial, J.: Applications of Combinatorial Optimization for Online Shopping. NAKOM, Poznan (2012)
18. Buyya, R., Broberg, J., Goscinski, A. M.: Cloud Computing Principles and Paradigms. Wiley Publishing (2011)
19. Foster, I., Zhao, Y., Lu, S.: Cloud Computing and Grid Computing 360-Degree Compared. In: Grid Computing Environments Workshop, 2008. IEEE (2008) 1–10
20. Zhang, Q., Cheng, L., Boutaba, R.: Cloud computing: state-of-the-art and research challenges. Journal of Internet Services and Applications 1 (2010) 7–18
21. Rimal, B., Choi, E., Lumb, I.: A taxonomy and survey of cloud computing systems. In: 5th International Joint Conference on INC, IMS and IDC. (2009) 44–51
22. Tantar, A., Nguyen, A., Bouvry, P., Dorronsoro, B., Talbi, E.: Computational intelligence for cloud management current trends and opportunities. In: IEEE Congress on Evolutionary Computation (CEC), IEEE (2013) 1286–1293
23. Kliazovicz, D., Pecero, J., Tchernykh, A., Bouvry, P., Khan, S., Zomaya, A.: Ca-dag: Communication-aware directed acyclic graphs for modeling cloud computing aplications. In: IEEE 6th International Conference on Cloud Computing, IEEE-Cloud 2013. (2013)
24. Chaisiri, S., Lee, B., Niyato, D.: Robust cloud resource provisioning for cloud computing environments. In: IEEE International Conference on Service-Oriented Computing and Applications (SOCA). (2010) 1–8
25. Castro, H., Villamizar, M., Sotelo, G., Diaz, C., Pecero, J., Bouvry, P.: Green flexible opportunistic computing with task consolidation and virtualization. 16 (2013) 545–557
26. Guzek, M., Varrette, S., Plugaru, V., Pecero, J., Bouvry, P.: A Holistic Model of the Performance and the Energy-Efficiency of Hypervisors in an HPC Environment. In Pierson, J.M., Da Costa, G., Dittmann, L., eds.: Energy Efficiency in Large Scale Distributed Systems. Lecture Notes in Computer Science. Springer Berlin Heidelberg (2013) 133–152
27. Pal, R., Hui, P.: Economic models for cloud service markets: Pricing and Capacity planning. Theoretical Computer Science 496 (2013) 113–124
28. Grozev, N., Buyya, R.: Inter-cloud architectures and application brokering: taxonomy and survey. Software: Practice and Experience 44 (2014) 369–390
29. Spillner, J., Brito, A., Brasileiro, F., Schill, A.: A Highly-Virtualising Cloud Resource Broker. In: IEEE/ACM Fifth International Conference on Utility and Cloud Computing. IEEE Computer Society (2012) 233–234

30. Nesmachnow, S., S., I., Dorronsoro, B., Talbi, E. G., Bouvry, P.: List scheduling heuristics for virtual machine mapping in cloud systems. In: VI Latin American Symposium on High Performance Computing (HPCLatAm). (2013) 1–12
31. Nesmachnow, S., S., I., Dorronsoro, B., Talbi, E. G., Bouvry, P.: A parallel hybrid evolutionary algorithm for the optimization of broker virtual machines sublet in cloud systems. In: 2nd International Workshop on Soft Computing Techniques in Cluster and Grid Computing Systems (SCCG). (2013) 1–12
32. Usha, M., Akilandeswari, J., Fiaz, A. S.: An efficient qos framework for cloud brokerage services. In: International Symposium on Cloud and Services Computing (ISCOS). (2012) 76–79
33. Carpentier, J., Gelas, J., Lefevre, L., Morel, M., Mornard, O., Laisne, J.: CompatibleOne: Designing an Energy Efficient Open Source Cloud Broker. In: Second International Conference on Cloud and Green Computing (CGC). (2012) 199–205
34. Nesmachnow, S., Dorronsoro, B., Pecero, J., Bouvry, P.: Energy-Aware Scheduling on Multicore Heterogeneous Grid Computing Systems. Journal of Grid Computing 11 (2013) 653–680
35. Iturriaga, S., Nesmachnow, S., Dorronsoro, B., Bouvry, P.: Energy efficient scheduling in heterogeneous systems with a parallel multiobjective local search. Computing and Informatics Journal 32 (2013) 1001–1022
36. Pinel, F., Dorronsoro, B., Pecero, J., Bouvry, P., Khan, S.: A two-phase heuristic for the energy-efficient scheduling of independent tasks on computational grids. Cluster Computing 16 (2013) 421–433
37. Alba, E., Dorronsoro, B.: Cellular Genetic Algorithms. Volume 42 of Operations Research/Computer Science Interfaces Series. Springer-Verlag US (2008)
38. Tantar, A. A., Danoy, G., Bouvry, P., Khan, S. U.: Energy-Efficient Computing Using Agent-Based Multi-objective Dynamic Optimization. In Kim, J.H., Lee, M.J., eds.: Green IT: Technologies and Applications. Springer Berlin Heidelberg (2011) 267–287
39. Dorronsoro, B., Danoy, G., Nebro, A. J., Bouvry, P.: Achieving super-linear performance in parallel multi-objective evolutionary algorithms by means of cooperative coevolution. Computers & Operations Research 40 (2013) 1552–1563
40. Nebro, A., Durillo, J., Luna, F., Dorronsoro, B., Alba, E.: MOCell: A Cellular Genetic Algorithm for Multiobjective Optimization. International Journal of Intelligent System 24 (2009) 726–746
41. Nebro, A., Luna, F., Alba, E., Dorronsoro, B., Durillo, J., Beham, A.: AbYSS: Adapting Scatter Search to Multiobjective Optimization. IEEE Transactions on Evolutionary Computation 12 (2008) 439–457
42. Blazewicz, J., Ecker, K. H., Pesch, E., Schmidt, G., Weglarz, J.: Handbook on Scheduling: From Theory to Applications. International Handbooks on Information Systems. Springer Science & Business Media (2007)
43. Pinel, F., Dorronsoro, B., Bouvry, P.: Solving very large instances of the scheduling of independent tasks problem on the gpu. Journal of Parallel and Distributed Computing 73 (2013) 101–110
44. Pecero, J., Huacuja, H., Bouvry, P., Pineda, A., Loces, M., Barbosa, J.: On the energy optimization for precedence constrained applications using local search algorithms. In: International Conference on High Performance Computing and Simulation (HPCS). (2012) 133–139