Multi-objective Scheduling for Distributed Security Services

J. A. Larco, U. Kaymak and R. Dekker
Econometric Institute, Erasmus University Rotterdam
P. O. Box 1738, 3000 DR, Rotterdam, The Netherlands
Phone: +31 – 10 – 408 1350, Fax: +31 – 10 – 408 9162
email: {kaymak, rdekker}@few.eur.nl

ABSTRACT: In this paper we deal with a common problem found in the operations of security and preventive/corrective maintenance services: that of routing a number of mobile resources to serve foreseen and unforeseen tasks during a shift. We define the Mobile Re-Allocation Problem (MRAP) as the problem of maximizing the weighted number of tasks on time. For solving the MRAP, we propose to solve successively a multi-objective optimization problem called the stochastic Team Orienteering Problem with Multiple Time Windows (s-TOP-MTW) so as to consider information about foreseen tasks and the arrival process of unforeseen tasks. Solving successively the s-TOP-MTW we find that there is in fact a trade-off between both objectives and that considering the reliability of routes may aid in maximizing the expected proportion of tasks accomplished on time. We also discuss ways of solving larger instances with the genetic algorithm NSGA-II.

KEYWORDS: Robust scheduling, security services, reliability, TOP, MEXCLP

1. INTRODUCTION

The domain of geographically distributed security services and preventive/corrective maintenance services involves the task to route mobile resources and/or vehicles over a number of sites so that various services can be performed. Such systems are usually characterized by the presence of multiple parties including teams of security personnel surveying an area autonomously, the operational management of the service company, the recipients of the security services and other stake holders. At an abstract level, the system can be considered as an ad hoc network in which the network nodes (different security teams, management centre and customers) communicate at (ir)regular times to exchange information and arrive at optimal decisions regarding allocation of tasks and routing of resources. Since alarms and other emergency occur during the course of operations, planning in this environment must necessarily be dynamic and real-time.

From a managerial perspective, the routing of mobile resources must remain cost effective while providing a high quality of service to the customers. One of the main dimensions of quality of service is that of the fulfilment of deadlines for performing each task, often specified by contracts. Satisfying the deadlines, while remaining within budget constraints, is a main concern for operational managers in the daily operations of distributed services. However, such objective is not trivial given the complexity of routing and re-routing mobile resources in non-deterministic environments where several sources of uncertainty exist such as unexpected delays due to heavy traffic, service vehicle breakdowns and the arrival of unforeseen, new tasks among others. In this paper we focus in what is considered by the security services the main source of uncertainty: the arrival of unexpected new tasks. The importance of unexpected emergency tasks is critical as it involves events such as a robbery in a security services context or people trapped in an elevator in a maintenance services context. Nonetheless, when evaluating the fulfilment of deadlines of a working-day shift, both types of tasks must be taken into consideration with their respective importance weight.

In this context, we define the Mobile Re-allocation Problem (MRAP) as the problem of devising a strategy (or decision procedure) for the allocation and re-allocation of mobile resources each time an unforeseen task arrives to maximize the expected weighted number of tasks[1] served within time windows and before the end of a shift. The time windows are defined as the earliest and latest allowable times to start servicing the tasks at their respective sites given that the clients' main concern is that their requests are being dealt with. In the case of foreseen tasks, such deadlines are specified on advance at the request of the client whereas in the case of unforeseen tasks, the deadlines are defined by a given response-time tolerance (i.e. mobile resources must arrive on site within a certain response-time from the time the alarm goes off). If it is known that the deadlines are not going to be met, the tasks are not served and back-up response is required from external parties (e,g. police).

[1] Including foreseen and unforeseen tasks.

As there is an interaction between the selected strategy to generate new routes and the arrival of new unforeseen events, it is intractable to calculate a-priori the expected weighted number of tasks served on time, let alone optimize such a number. As a result, it is common practice to solve successively a combinatorial problem that assumes complete information like a Travelling Salesman Problem (TSP) variant for the MRAP.

The consequence of such approach is that it disregards information about the possible future arrivals of unforeseen tasks. In this paper we propose as the strategy to address the MRAP to solve successively a multi-objective problem that we refer to as the stochastic-Team Orienteering Problem [1] with Multiple Time Windows (s-TOP-MTW). The first objective is an Orienteering Problem objective variant that maximizes the foreseen weighted number of tasks served on time. The second objective is a variant of the Maximum Expected Coverage Location Problem (MEXCLP) and consists of maximizing a reliability measure defined as the expected weighted number of unforeseen requests handled assuming no modifications of routes.

To test the effectiveness of this proposal the paper is structured as follows: In Section 2 we develop the first objective of the s-TOP-MTW dealing with foreseen tasks based on the Orienteering Problem (OP). The second objective of the s-TOP-MTW is presented in Section 3 dealing with information about the arrival of new unforeseen tasks based on the Maximum Expected Location Problem (MEXCLP). Next, in Section 4 we discuss some details of the MOP integration of these objectives and provide a simple algorithm to solve it in a dynamic setting. We then test in Section 5 the suitability of this proposal in a series of experiments. In Section 6 we conclude our exposition of how risk considerations may enhance deadline fulfilment and how large scale problems can be solved with the aid of nature-inspired solving techniques.

2. THE ORIENTEERING PROBLEM AND VARIANTS

In sharp contrast to the TSP, the OP does not deal with finding a path that minimizes the total distance travelled to visit all sites distributed in a network; the OP in fact does not even require visiting all the sites. The OP's purpose deals instead with maximizing the total rewards associated to visiting each site within a total distance budget, $c_{\max}$ [1].

When the OP is modified to include instead of a single path multiple paths travelled by multiple resources then the problem transforms into the Team Orienteering Problem or TOP as defined in I-Chao et al. [2]. If we further include the possibility of several tasks per site and starting time windows associated with each task denoted by (e_n, l_n) then the resulting problem would be the Team Orienteering Problem with Multiple time Windows or TOP-MTW. Solving the TOP-MTW would in effect maximize the "expected"[2] weighted (weighted by a relative importance factor w_n) number of tasks served on time if the MRAP would have complete information with all tasks known in advance at the start of the shift. If $|K|$ guards were to serve the requests and if x_{kn} indicates whether a guard $k \in K$ serves task $n \in T$ then the following would result in the TOP-MTW* normalized objective:

$$z_{OP*} = \frac{1}{\sum_{n \in T} w_n} \sum_{k \in K} \sum_{n \in T} w_n x_{kn} \tag{1}$$

If all the tasks considered in the TOP-MTW are served, then the equivalent of the TSP variant, the multi Travelling Salesman Problem with Multiple Time Windows or m-TSP-MTW (see [3] for a version with single time windows) would have a feasible solution space. Thus, in such a case, the optimal solution of the m-TSP-MTW will be also optimal for the TOP-MTW. However, it is by no means guaranteed that in a real setting all the tasks could be served due to conflicts between time windows and/or the shift length as a constraint. Thus, for security service applications the TOP-MTW is preferable over the m-TSP-MTW.

However, the main question that remains is whether solving successively a TOP-MTW, disregarding at the same time information of the stochastic arrival process of new tasks can be enough to obtain, good, near-optimal solutions for the MRAP. To provide insight into this question consider the next example where there is a fully connected network and

[2] The word "expected" in this case is irrelevant as the problem has complete information.

there are six sites with six tasks to be served and one central base where two guards (mobile resources) must start and end their shifts from 0 hrs to 10 hrs:

Task index / {Site Location}	1{a}	2{b}	3{c}	4{d}	5{e}	6{f}
X-coordinate (km)	-1.0	1.0	-1.0	1.0	-1.0	1.0
Y-coordinate (km)	1.0	1.0	0.0	0.0	-1.0	-1.0
Earliest allowable starting time(hr)	2.5	1.2	1.0	3.0	1.2	2.5
Latest allowable starting time (hr)	5.0	7.0	9.5	6.5	1.5	5.0
Relative Importance of foreseen / {unforeseen} tasks (weight)	16.7 {100.0}					
Duration of foreseen / {unforeseen} tasks (hr)	1.0 {1.5}					
Poisson Distributed Mean rate of unforeseen tasks (# tasks)	1 (same chance for all sites of occurrence)					
Speed of resources (km/hr)	1					
Response Time Radius (hr)	0.707					

Table I:. Example of a security system

At the start of the shift, we can solve for TOP-MTW and obtain alternative optima. One of such alternative optima yields a solution set A where guard 1 visits the sites in the sequence {a, b, c} while guard 2 visits the sites in the sequence {e, f, d}. Another of such alternative optima yields a solution set B where guard's 2 visiting sequence remains the same while guard 1 visits the sites in an anti-clockwise manner in the sequence {c, a, b}. Both solutions collect a relative importance weight of 100, however, are these two solutions equally preferable in a dynamic situation with emergency tasks? To answer this question we illustrate a snapshot at time 7.50 hrs showing the areas that can be served within standard response time (i.e. the areas covered) by each guard, drawing a radius of 0.707 with each guard at its centre:

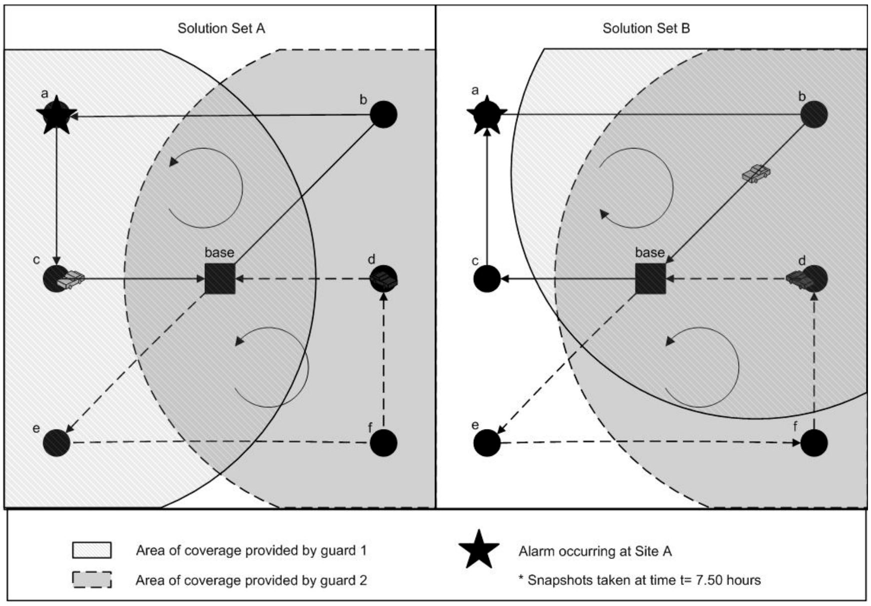

Figure 1: Coverage considerations do matter

As seen in the figure, in Solution Set A, the guards travel both in the same anti-clockwise way and thus are always located in opposite ends and hence cover all the sites. However, in Solution Set B, Figure 1 shows that the guards travel in different direction (i.e. guard 1 travels clockwise while guard 2 travels anticlockwise) and are at the same side of the figure, either left or right implying that certain sites {a, b, c} are not covered as at time 7.50 hours. If an alarm occurs at site {a} at this moment in time, guard 1 in Solution Set A would be able to attend it while neither guard in Solution Set B would. Hence, if such a pattern is observed throughout time then it is logical to select Solution Set A over Solution Set B based on the fact that is less risky. Nonetheless, it is still important to quantify the level of risk that a certain route implies, particularly if there may be trade-offs between risk and weighted number of known tasks. In the next section, we provide new insight as to how to measure such a risk.

3. THE MAXIMUM EXPECTED COVERAGE LOCATION PROBLEM

Among the models in the literature of urban planning location referred as emergency sitting models (see [4] for a review of emergency sitting models), the MEXCLP is particularly relevant for the MRAP. The MEXCLP's purpose is to locate M emergency vehicles distributed over a number of eligible locations $j \in J$ so as to be able to respond on time when an emergency request occurs in any of the sites $i \in I$. The MEXCLP is relevant as it takes into account the idea of risk inherent to the location of resources which was taken into account in the OP variants in the previous section.

Given that q is the busy fraction or the probability that a resource is not available to serve a task for being busy serving another one, w_i is the associated weight importance of each task generated at site $i \in I$ and y_{ik} is a decision variable that indicates if at least *k* resources are covering site $i \in I$, the objective function of the MEXCLP is as follows:

$$z_{MEXCLP} = \sum_{i \in I} \sum_{k=1}^{M} (1-q)(q)^{k-1} w_i y_{ik} \qquad (2)$$

An easy way of interpretation is to make all the weights w_i equal to 1, and then the MEXCLP calculates in effect the probability that if an unforeseen task arrives there would be available at least one resource to serve it. The probability that the demand at site $i \in I$ will be assigned to its *k*-th-coverer is $(1-q)(q)^{k-1}$ as it is the probability that the $k-1$ previous coverers have not covered it due to being busy (i.e. q^{k-1}) multiplied by the probability that the k -th coverer is available with probability $1-q$.

The difference with the MEXCLP and the problem dealt in this paper is that the emergency vehicles in the MEXCLP serve from a fixed location, whereas in the MRAP tasks are served consecutively one after the other without needing to return to the base until the end of the shift. As a result, to adapt the MEXCLP to the purposes of the security services re-allocation problem the mobility of the resources should be considered. Thus, we propose an average normalized reliability measure over a shift that can be obtained by subdividing the shift interval $(0, TT)$ in N intervals, where snapshots are taken showing the updated locations of mobile resources:

$$z_{MEXCLP*} = \frac{1}{N \sum_{i \in I} a_i} \sum_{n=1}^{N} \sum_{i \in I} \sum_{k=1}^{M} (1-q_k)(q_k)^{k-1} a_i y_{ikn} \qquad (3)$$

It must be noted also that in the adapted MEXCLP*, the busy fraction q_k is no longer system wide and it is calculated individually per mobile resource sorted in descending order. The busy fraction is the ratio of the expected workload of the mobile resource[3] and the available time left to serve requests. The busy fraction also takes into account that guards are only considered available when the tasks are either en-route to serve an ordinary task or waiting at a site for serving an ordinary task before the earliest allowable time.

4. DYNAMICALLY SOLVING THE MULTI-OBJECTIVE S-TOP-MTW

The information available for routing the mobile resources can be classified in two categories: deterministic and probabilistic. While the TOP-MTW* can deal with deterministic information about known tasks to be performed, a modified MEXCLP* can deal with probabilistic information about the arrival process of new unforeseen tasks. Both objectives are normalized to have comparative scales from 0-1, provided that for scaling purposes the total expected weight of unforeseen tasks[4] is equal to the expected and the total weight of foreseen task are the same in the MRAP. If the weighted total expected weighted is not the same to the weight of foreseen tasks such unbalance should be factored in. In any case, it is not trivial to know how to balance these objectives even if they are normalized, hence the need for solving a multi objective problem.

[3] Workload generated by foreseen tasks and by unforeseen tasks covered by the resource.
[4] Calculated multiplying the expected number of unforeseen tasks in a shift with the deterministic weight of unforeseen tasks.

Since only one solution must be selected each time an alarm occurs, an affine combination of objectives in (1) and (3) is used such that the s-TOP-MTW's objective is $z^* = (1-\alpha)z_{TOP-MTW^*} + \alpha z_{MEXCLP^*}$ providing arbitrary values of α to be tested as shown in Section 5. For the solution which consists of only a sequence of tasks (including foreseen tasks and unforeseen tasks that have arrived) executed by mobile resources, it is assumed that when the mobile resources finish their tasks they immediately travel to the next and wait until the earliest allowable starting time to serve the task. The basic dynamic application of the solution algorithm outline is as follows:

1. Start Loop A: Recalculate position of mobile resources, status and liberation times.
2. Start Loop B: Select randomly a mobile resource.
3. Start Loop C: Verify the feasibility of inserting a task in terms of fulfilling starting service time window and being able to return to base before end of shift.
4. Calculate fitting function: $f(n) = \left\{ \frac{w_n}{[\max(Last_k, e_n) + d_n] - Last_k} \right\}^4$ where w_n is the associated importance (weight) of task $n \in T$, $Last_k$ is the finishing serving time of the last task scheduled in mobile resource *k*-th route, e_n is the earliest allowable starting time of the task and d_n is the duration of the task.
5. Store the four best feasible tasks to be added. End Loop C.
6. Select stochastically feasible task, with probability dependent on fitting function $f(n)$. End Loop B.
7. Track route positions of mobile resources and calculate average reliability measure: MEXCLP*.
8. Select non-dominated solution with a certain α integration factor.
9. Generate new alarm until shift finishes. End Loop A.

In Loop A, the actual simulation is carried out by generating new unforeseen tasks and updating the new information set to be optimized observing the new location of mobile resources and calculating their liberation times given that for example a guard serving an emergency task can not be interrupted (i.e. as mentioned earlier, only guards en-route or waiting for starting time windows in ordinary tasks can be interrupted).

In Loops B and C an optimization algorithm is executed based on that of Tsiligirides' stochastic algorithm [5]. The algorithm is a construction algorithm that adds tasks to a route stochastically with probabilities dependant on a fitting function which is the ratio of the associated importance weight of the task and the additional time spent for serving such a task. The extension to Tsiligirides' algorithm consists in the fact that more than one mobile resource is considered and thus, mobile resources are randomly chosen for insertion of a single task at a time. Also Tsiligirides' algorithm is extended in the sense that it considers the starting time windows influence in the extra time used, making it unlikely to select a task whose earliest allowable starting time is yet distant from the last finishing serving time, $Last_k$.

5. EXPERIMENTS

All the experiments presented in this paper have a common fully connected network of 10 sites[5] with starting service time windows that on average are 2/3 the length of the whole time shift length that is 40 time units. The number of unforeseen tasks in a shift is Poisson distributed and has a mean of 2 tasks. There is a heterogeneous frequency of alarms, 2 sites account for 50% of the unforeseen tasks. The duration of emergency tasks is 1.25 time units and the duration of ordinary tasks varies form from 0.37 to 1.48 time units.

The first experiment, using a response tolerance time of 7 time units, consists of determining if there is effectively a trade-off between the normalized TOP-MTW* and MEXCLP* by constructing an efficient frontier with all the identified non-dominated solutions. The results obtained are as follows:

[5] In preliminary runs, the algorithm only proved effective in finding "good" solutions for small instances of the problem (maximum 12 sites with up to 2 tasks per site).

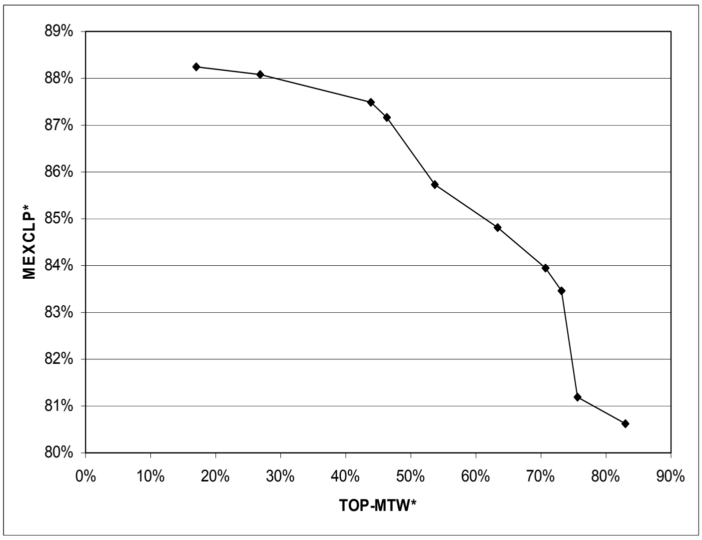

Figure 2: Experiment 1: MEXCLP* vs. TOP-MTW* trade-off

The experiment verified that a trade-off does exist (i.e. a reduction of 8% of MEXCLP* in exchange for an increase in 65% in TOP-MTW*) and hence a hierarchical optimization proves insufficient for solving successively the re-allocation of guards. Therefore we need a multi-objective optimization approach.

The second experiment consists of investigating the sensitivity of the reliability measure MEXCLP* to variations in the response tolerance parameter, testing also three different values of α , to observe how important the adjusting of the response time can be for the reliability of routes. Figure 3 show a characteristic "S" shape from the results obtained:

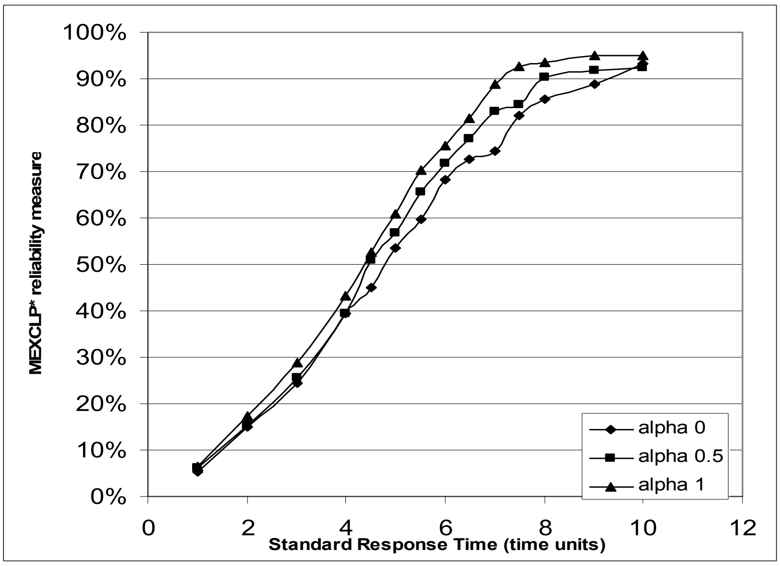

Figure 3: Experiment 2: Sensitivity of MEXCLP* reliability measure to standard response time

It is interesting to note from Figure 3 that there is a very sensitive region to small variations of standard response time (in this case around the value 7 time units). The managerial implication of this sensitivity is that by investigating the impact of standard response times on reliability, managers may renegotiate slightly higher response time parameters to offer significantly superior reliability.

The third experiment is the most important of all as it tries to answer the question as how to balance both objectives to obtain the maximum expected weighted number of tasks accomplished on time. A series of 60 simulation runs were carried out, applying successively the algorithm described in Section 4 with different values of α . The results are expressed in % of possible weighted number of tasks accomplished on time during the day-shift simulation run. They are as follows:

α=	0.00	0.000001	0.25	0.50	0.75	0.999999	1.00	In hindsight
Average of % of weighed tasks accomplished in due time	63.85	64.98	65.01	63.75	59.72	35.18	39.65	72.68
Average of mean coverage %	61.18	63.75	64.27	68.26	69.93	72.96	72.85	62.88

Table II:. Experiment 3: Simulation run results.

The results show that using the MEXCLP* can positively influence the expected % of weighed tasks accomplished in due time. However, this is only for low values of α, indicating that reliability is of second importance after that of greedily maximizing the weighed amount of known tasks accomplished in time. Moreover, for high values of α the results are poor and this can be explained since although there is a high percentage of coverage, known tasks, including emergency tasks released are not considered as important and thus it is never time to "cash-in" the good positioning.

6. CONCLUSIONS AND FURTHER RESEARCH

The results of the third experiment prove that there is a degree of short-sightedness of only solving a regular routing model such as the TOP-MTW*, indicating that considering also the risks at a secondary level by measuring the inherent risks of routes through the MEXCLP* objective in the s-TOP-MTW may prove profitable. However, the results also clearly point out that weighting the reliability measure more than the TOP-MTW* score may prove disastrous.

On the other hand, the results call for more research so as to observe if these findings are sustained in other layouts of sites and increased relative workload due to emergency tasks. Moreover, for enhancing the applicability of the model, larger-scale problems have to be solved. Given that a multi-objective solving mechanism is needed and that the interplay between reliability and tasks services is poorly understood[6] the NSGA-II [7] may prove adequate, particularly due to its "low" computational complexity $o[MN^2]$.

In the implementation of the TOP-MTW, a similar three-phase crossover operator as suggested for the similar TSP with Profit problem may be applied [8]. Such method consists in first modifying the parents by eliminating all nodes that are not common to both parents, next apply an edge recombination crossover (ERX) and finally the sub-paths that involve the nodes discarded in the first phase are added to the offspring. For the mutation, three operators are also suggested namely, 1) insert/remove node, 2) swap nodes, 3) local neighbourhood search.

However, it must be noted that although traditional multi-objective algorithms attempt to evenly distribute non-dominated solutions so as to sample the greatest range possible of the efficient frontier, given the initial results, this might not be needed for the present application. In fact, it may be useful to concentrate on a small region of the efficient frontier that yield better results in the expected % of weighed tasks accomplished on due time. Thus, the crowd comparison operator of the NSGA-II [7] may still be applied but for a limited range (a range of α may be used for this purpose).

Important extensions worth investigating are investigating the effect of free delays and having alternative paths to transverse between tasks as it is possible to obtain important improvements in reliability of routes.

REFERENCES

[1] Golden, B; Assad, A; and Dahl, R, 1984, "Analysis of a large-scale vehicle routing problem with an inventory component", Large Scale Systems 7, pp. 181-190.

[6] Moreover, the TOP-MTW* is a variant of the OP which is known to be NP-hard [6].

[2] I-Ming Chao; Golden, B; Wasil, E, 1996, “The Team Orienteering Problem”, European Journal of Operational Research, 88 pp. 464-474.

[3] Mitrović-Minić, S; Ramesh K, 2006, “The multiple TSP with time windows: vehicle bounds based on precedence graphs” Short Communication Operations Research Letters, 34, pp. 11-120

[4] Revelle, C, 1989, “Review, extension and prediction in emergency service sitting models”, European Journal of Operational Research, 40 pp. 58-69.

[5] Tsiligirides, T, 1984, “Heuristic Methods Applied to Orienteering”, The Journal of Operations Research Society, 35 pp. 797-809.

[6] Golden, B; Levy, L; and Vohra, R, 1987, “The orienteering problem”, Naval Research Logistics, 34, pp. 307-318.

[7] Deb, K; Pratap, A; Agarwal, S; Meyarivan, T, 2002, “Fast and Elitist Multi Objective Agorithm: NSGA-II” 2002. IEEE transactions on evolutionary computation, 6 pp. 182-197.

[8] Jozefowiez, N; Laguna, M; Glover; F, 2006, “A Multi-Objective Evolutionary Algorithm for the Bi-Objective Travelling Salesman Problem with Profits”, MOPGP’06: 7th International Conference of Multi-objective Programming and Goal Programming.