Internet Prospect's flow forecasting for a multi-period optimization model of offer/Demand assignment problem.

Manel Mammar, Vincent Mousseau, Wassila Ouerdane, Alexandre Aubry

HAL Id: hal-01205141
https://hal.science/hal-01205141

Submitted on 26 Sep 2015

INTERNET PROSPECTS' FLOWS FORECASTING FOR A MULTI-PERIOD OPTIMIZATION MODEL OF OFFER/DEMAND ASSIGNMENT PROBLEM

Manel Maamar[1], Vincent Mousseau[1], Wassila Ouerdane[1] and Alexandre Aubry[2]

[1]LGI, CentraleSupelec, Grande Voie des Vignes, 92295
{manel.maamar,vincent.mousseau,wassila.ouerdane}@ecp.fr

[2] Place des Leads, 15 bis boulevard Saint-Denis, 75002 Paris, France
alexandre.aubry@placedesleads.com

ABSTRACT

In this paper, we propose a forecasting system for a multi-period assignment model, where Internet prospects' flows evolve continuously over time. In particular, we show how we build a forecasting model of Internet prospects' flows while considering some characteristics of a multi-period assignment model in order to optimize and improve the marketplace's system of Place des leads (PdL). A numerical investigation is conducted based on real data of PdL application to show the benefits of introducing forecasting information to the multi-period optimization model.

Keywords: Assignment problem, multi-period optimization, Internet prospects forecasting

1   INTRODUCTION

Various methods of forecasting have been proposed in the literature. The aim of such a methods is to reduce the uncertainty related to the unknowing of the future. Several approaches of forecasting, namely endogenous, exogenous and hybrid exist (see [2]). Most of these methods have been implemented and regrouped in *R* software [6]. On the other hand, a well-known problem is the standard Assignment Problem (AP) where the aim is to optimally resolve the problem of assigning n jobs (tasks) to n individuals (machines), such that minimum cost or maximum profit can be obtained [3]. The classical AP has a wide range of applications that span manufacturing, personnel scheduling and transportation, to name a few [4, 8]. Various algorithms were developed to find solutions efficiently [1, 5]. For a standard AP, only the cost or profit from each possible assignment is considered in the problem formulation. However, in real applications other issues can be addressed.

In this paper, we study a variant of the AP that involves assigning Internet prospects to orders over several time periods using two objectives to guide the search for solutions. The AP concerns Place des Leads (PdL), a marketplace where several clients (assurances, banks...) make orders of Internet prospects. These later are defined in a Web context as a contact ready to carry out an action (purchase, estimate...) [7]. More precisely, we aim to combine a forecasting model with an optimization model in order to enhance the PdL's assignment problem results. The majority of the research works deals with either forecasting models or optimization models. In this work, our idea is to integrate forecasting model into a multi-period assignment in order to improve assignment by integrating knowledge about Internet prospects' flows continuously evolving in the optimization model.

To account for that, we introduce, in section 2, the multi-period assignment model and its characteristics. In section 3, (i) we discuss the relationship between both forecasting model and optimization one, (ii) we propose a forecasting system which considers all characteristics of the multi-period optimization model and (iii) we show how this forecasting is integrated in the multi-

period optimization model. To show the efficiency of the proposed model, numerical tests are addressed in section 4. A conclusion and some perspectives are discussed in section 5.

2 MULTI-PERIOD OPTIMIZATION MODEL OF OFFER/DEMAND ASSIGNMENT

In this section, we introduce a multi-period assignment model, defined in [7], where the aim is to establish an adequacy of offer and demand with consideration of several periods. In particular, the objective is to optimize turnover while: (i) considering instantly new offer and demand of Internet prospects and (ii) taking into account the eventual future system updates. The idea is to optimize assignment, at each time step, by leveraging knowledge about evolution of Internet prospects' flows as well as updating of orders' quotas, over several future periods.

The multi-period assignment problem is characterized by moving horizon, which means that time is modeled through several time steps. At each time step t, we consider a set T of periods such that $T = \{t': t' \in \{1, .., g\}\}$, where $t' = 1$ is the current period and $t' \in \{2, .., g\}$ are future periods. Thus, at each time step t, we have to decide which Internet prospects will be assigned to which orders, while considering all time periods of the set T. We assume that both orders and clients are the same over all periods t'. However, it is possible to have a new Internet prospects' flows arriving in an indeterminist way over all future periods t'. Consequently, at each period t', we have $m_{t'}$ Internet prospects and n orders coming from k clients that can make one or more orders, represented respectively by the following sets:

- Set of Internet prospects: $P = \{j: j \in \{1, .., m_t\}\}$.
- Set of orders: $O = \{i: i \in \{1, .., n\}\}$.
- Set of clients: $C = \{k: k \in \{1, .., r\}\}$.

The model P represents a deterministic linear mathematical program which optimizes, at each time step t, Internet prospects' assignment to orders by considering several future periods, whose components are presented in what follows.

$$
P \begin{cases}
\max \quad f = \sum_{t'=1}^{g} \sum_{i,j:x_{ijt'} \in X_{t'}} P_i \ x_{ijt'} \\
s.t. \\
\sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, & \forall i \in O, \ \forall t' \in T \quad (1) \\
\sum_{i,j:x_{ijt'} \in X'_t} a_{ik} \ x_{ijt'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ijt''} \in X_{t''}} a_{ik} \ x_{ijt''}, & \forall k \in C, \ \forall t' \in T \quad (2) \\
S_j + \sum_{t''=1}^{t'-1} \sum_{i:x_{ijt''} \in X_{t''}} x_{ijt''} + b \sum_{i:x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i:x_{ijt'} \in \overline{E_{t'}}} x_{ijt} \leq b, & \forall j \in P, \ \forall t' \in T \quad (3) \\
\sum_{t'=1}^{g} \sum_{i:x_{ijt'} \in X_{t'}} a_{ik} \ x_{ijt'} \leq 1, & \forall j \in P, \ \forall k \in C \quad (4) \\
x_{ijt'} \in \{0, 1\}, & \forall i \in O, \ \forall j \in P \ \forall t' \in T
\end{cases}
$$

At each period t', Internet prospect's assignment to order is represented by the following binary variable:

$$
x_{ijt'} = \begin{cases} 1 \ if \ Internet \ prospect \ j \ is \ assigned \ to \ order \ i \ at \ period \ t' \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad otherwise \end{cases}
$$

Moreover, the set of all existing "order/Internet prospect" assignments at period t' is described by the set $X_{t'} = \{x_{ijt'} : i \in O, j \in P, t' \in T\}$. However, in each $X_{t'}$, some pairs $x_{ijt'}$ are not possible. Thus, we will consider that $x_{ijt'} \in X_{t'}$ only if an Internet prospect j corresponds to an order i, at a period t'. Moreover, some orders ask for Internet prospects in exclusivity. It means that Internet prospect is sold at most once. Otherwise, we can sell Internet prospect several times to different orders. We call this sale: non exclusivity. Accordingly, we have $E_{t'}$: set of $x_{ijt'}$, at period t', such that i is an exclusive order and $\bar{E}_{t'}$: set of $x_{ijt'}$, at period t', such that i is a non exclusive one. Thus, $X_{t'} = \{E_{t'} \cup \bar{E}_{t'}\}$ such that $t' \in T$.

Internet prospects' assignment problem while considering, at each time step t, g periods t' is submitted to the followings constraints.

- Constraint (1) (Constraint (2)) makes sure that the number of Internet prospects assigned to each order i (client k respectively), at each period t', respects remaining volume of the order, which is the difference between remaining volume of order i (client k respectively) at period 1, and delivered Internet prospects' number to order i (client k respectively) from period 1 to happening period t'. Such that:
  - Ro_i (Rc_k) is Internet prospects' number remaining to be delivered to order i (client k respectively), at period 1, which is defined as: $Ro_i = min\ \{Go_i, Ho_i, Do_i, Mo_i\}$ ($Rc_k = min\ \{Hc_k, Dc_k, Mc_k\}$ respectively), where: Go_i is global volume of order i; Ho_i (Hc_k) is hourly quota of order i (client k respectively); Do_i (Dc_k) is daily quota of order i (client k respectively) and Mo_i (Mc_k) is monthly quota of order i (client k respectively).
  - a_{ik} is a binary variable described as follow:

$$a_{jk} = \begin{cases} 1 & \text{if Internet prospect } j \text{ is assigned to client } k \\ 0 & \text{otherwise} \end{cases}$$

  - Constraint (3) expresses the fact that, over all g periods considered in the optimization, each exclusive Internet prospect can be sold no more than once and each non exclusive one can be sold no more than b times. Such that: S_j is number of previous sales before period 1 of Internet prospect j, and b is sales' number of non exclusive Internet prospects.
- Constraint (4) guarantees that each client k receives each Internet prospect no more than once over all g periods considered in the optimization.

Finally, function f represents the multi-period model's objective function which seeks to maximize turnover, over all periods considered in the optimization. Such that: P_i is sell price of order i.

3 INTERNET PROSPECTS' FLOWS FORECASTING

3.1 Why should we consider forecasts?

We are in a context where Internet prospects' flows evolve continuously in a determinist way over time. In order to optimize the turnover, we have to take into account this Internet prospects' flow in the assignment process. In other words, multi-period assignment model defined in the previous section allows optimizing assignment by leveraging knowledge about evolution of Internet prospects' flows. In the sense that, we provide to our model the set of all possible assignments of each period

t' considered in the optimization. Consequently, we have to compute Internet prospects' flows arriving over these periods.

Defining a forecasting model of Internet prospects' flows is therefore necessary to improve optimization model. Therefore, we propose in what follows to articulate a forecasting model with an optimization model.

3.2 Forecasting/Optimization articulation

Building a forecasting model requires to define which data to forecast before considering any forecasting method. In other words, we have to choose the appropriate historical data, time step and horizon which are the parameters of the forecasting model. The challenge is to set the parameters that will fit with the functioning of the optimization.

3.2.1 Data to forecast

In order to integrate Internet prospects' flows in the optimization model, we have to forecast at each period *t'* considered in the optimization model: (i) the number of Internet prospects arriving at this period and (ii) orders for which these Internet prospects can be assigned. A first approach could consist in forecasting the number of Internet prospects for each order. However, this idea is unsatisfactory because a large part of Internet prospects is shared by several orders. Consequently, if we forecast Internet prospects' number for each order, then Internet prospects shared by several orders will appear in the historical data of each of these orders. Thus, they will be counted several times. In addition, we need to know for each foreseen Internet prospect, the set of orders whose characteristics are satisfied. However, forecasting of Internet prospects' number for each order does not provide information about correspondence of these foreseen Internet prospects to the other orders.

To overcome the previous problems, we propose to forecast Internet prospects' number for each orders' group sharing the same Internet prospects. Accordingly, we will forecast Internet prospects' number as well as orders for which each foreseen Internet prospect can be assigned. Thus, each group of orders represents a time series for which a forecasting study will be conducted. How to build such order's group is discussed in section 3.3.1.

3.2.2 Forecasting time step and horizon

We face a complex forecasting problem, where the challenge is to integrate the forecasting results into a multi-period assignment model. The first difficulty comes from the fact that the forecasting of Internet prospects should be compatible with the multi-period optimization problem described in section 2. The second difficulty is related to the fact that the assignment of Internet prospects to orders is grounded, at each time step, on a rolling horizon of the future periods. Consequently, we have to compute Internet prospects' flows over these future periods such that, on the one hand, forecasting time step must be equal to the period of the multi-period assignment model, on the other hand, horizon should represent the number of the future periods considered in the multi-period assignment model.

The fact that forecasting time step and period of the optimization model (*t'*) must be the same generates another difficulty. In fact, the orders' quotas are updated at each hour. Therefore, integrating such information in the optimization model requires considering periods *t'* of one hour [7]. However, one hour forecasting time step does not necessarily guarantee a good forecasting because Internet prospects' flows density is low on this period. Therefore, we should consider periods *t'* greater than one hour in order to improve forecasting. The consequences of this, only some hourly updating of orders' quotas are integrated in optimization model. For example,

considering periods *t'* of 4 hours conducts to take into account only updating at 00 o'clock, 4 o'clock, 8 o'clock …. although we have a quotas' updating at each hour. However, forecasting with 4 hours time step will be better than forecasting with one hour time step because Internet prospects' density is more important with 4 hours time step. Accordingly, we have to choose a compromise solution.

3.3   Forecasting system

In the previous sections, we highlighted both the benefits and the difficulties of integrating forecasting into the optimization model. In this section, we present: (i) an algorithm which computes the time series and (ii) a forecasting model as well as its integration in the multi-period optimization model.

3.3.1 Time series' construction

The following algorithm computes the set of time series. More precisely, each time series, which is an orders' group sharing the same Internet prospects, provides the Internet prospects' number of this orders' group at each historical period.

```
Time_Series ():
Data:
```

- $T = \{t: t \in \{1,..., d\}\}$ is set of historical periods.
- $P = \{j: j \in \{1, ..., m_t\}\}$ is set of Internet prospects.
- $O = \{i: i \in \{1, .., n\}\}$ is set of orders.
- $\delta = \{l: l \in \{1, .., e\}\}$ is set of orders' groups.
- $A = \{y_{lt}: l \in \delta, t \in T\}$ such that y_{lt} is Internet prospects' number shared by orders' group l at period t.
- $X = \{x_{ijt}: i \in O, j \in P, t \in T\}$ such that $x_{ijt'} = 1$ if Internet prospect j corresponds to the characteristics of order i at period t, 0 otherwise.

```
Begin
 δ := ∅;
 For  t ∈ T  do {
    For   j ∈ P  do {
         l := Order_Group (j , t , O , X);
         if   l ∈ δ   then { y_lt := y_lt +1;}
         else {
                δ := δ ∪ {l};
                For   t' ∈ T   do {
                     if   t' = t   then {y_lt := 1;}
                     else {y_lt := 0;}
                }
           }
      }
   }
 return A;
End
```

The function **Time_Series()** generates the set of the time series. For each orders' group l, we have the time series $(y_{lt})_{t \in T}$ which is the Internet prospects' number of the orders' group l at each historical period t. The orders' group sharing the Internet prospect j at the period t is computed by the function **Order_Group(j , t , O , X).**

```
Order_Group (j , t , O , X):
```

```
Data:
    • l is an orders' group.
Begin
 l := ∅;
    For  i ∈ O  do {
         if  x_ijt = 1  then { l := l ∪ {i}};
         }
 return l;
End
```

3.3.2 Forecasting model and its integration in the multi-period optimization model

In order to compute the forecasting of each time series $(y_{lt})_{t \in T}$ such that $l \in \delta$, defined in the previous section, we use *ETS* forecasting model implemented in *R* which takes as input a time series and returns the appropriate smoothing method based on the time series. In our case, *ETS* forecasting model returns, at each period, the number of Internet prospects (IP) shared by the same orders' group.

To integrate Knowledge on new Internet prospects' flows in the *g* periods considered in the optimization model, we need both number of Internet prospects at each period as well as orders for which these Internet prospects can be assigned. For instance, if we consider the following example: first, the assignment is done on *3* periods, namely *0, 1 and 2*. Second, we have three orders and two Internet prospects at period *0*, such that the possible assignments are show in Table 1. Finally, the results of *ETS* forecasting are depicted in Table 2. Given these information, the question is how to build the set of all possible for period 1 and 2. We may integrate, at period 1, one IP for order 1 (see table 3) and two IPs, at period 2, for orders 2 and 3 (see table 4).

Table 1: Set of all possible assignments at period 0

order/IP	IP 1	IP 2
order 1	x_{110}	x_{120}
order 2	x_{210}	/
order 3	/	x_{320}

Table 2: *ETS* forecasting

orders' group/period	period 1	period 2
{1}	1	0
{2,3}	0	2

Table 3: Set of all possible assignments at period 1

order/IP	IP 1	IP 2	IP 3
order 1	x_{110}	x_{120}	x_{131}
order 2	x_{210}	/	/
order 3	/	x_{320}	/

Table 4: Set of all possible assignments at period 2

order/IP	IP 1	IP 2	IP 3	IP 4	IP 5
order 1	x_{111}	x_{121}	x_{131}	/	/
order 2	x_{211}	/	/	x_{242}	x_{252}
order 3	/	x_{321}	/	x_{342}	x_{352}

4   EXPERIMENTATION

In this section, we analyze the behavior of the proposed forecasting system. We perform a numerical investigation using real data from PdL database. Our aim is to check empirical validity of the proposed approach and validate its practical usefulness. Experiment is designed to address the

ability of the forecasting to optimize the results of the multi-period assignment model in terms of turnover. To do so, we perform assignment using both multi-period assignment model which includes an exact forecasting, i.e. the current data; we call this OF system and multi-period assignment model without forecasting, called O system. We compute the turnover for both systems OF and O, and evaluate the benefit induced by the introduction of an error free forecasting.
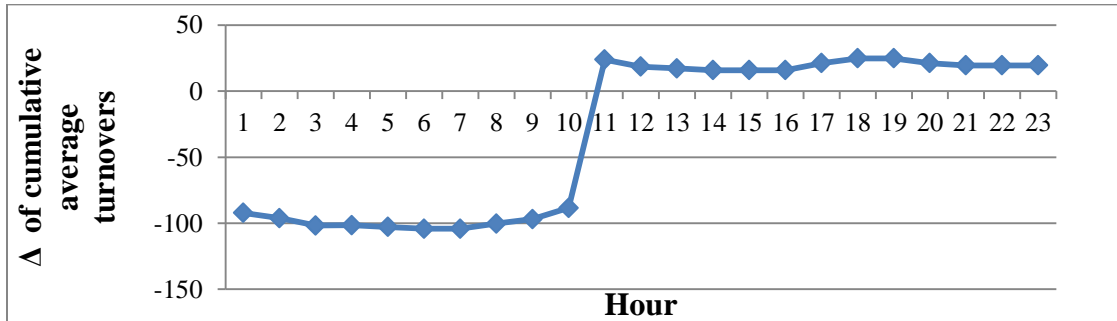

Figure 1: Difference between average turnovers of both OF system and O system over one day

Figure 1 represents the evolution, over one day, of the difference between the average turnovers of both OF system and O system. At the beginning of the day, the OF system generates an average turnover less than O system's average turnover. This situation continues until 6 o'clock where the OF system begins to be better than the O system. At 10 o'clock, the OF system retrieves all lost sales compared to O system and continues to be better until 10h30. After that, it begins to lose compared to O system. At the end of the day, OF system generates turnover more important than O system. This observation is not surprising because O system makes, at each current period, all possible assignments. However, the OF system may do not satisfy some orders, at the current period, in order to perform sales more expensive later. In fact, the system has information about the evolution of Internet prospects' flows over the future periods. We see this phenomenon at the beginning of the day as well as at the end. Clearly, in order to fully appreciate the gain of the OF system, we have to replay the experiment more than one day, an example is detailed in Table 5.

Table 5 Difference between average turnovers of both OF system and O system over three days

Day	Δ of cumulative average turnover
1	0
2	20
3	60

Table 5 shows the evolution, over three days, of the difference between the average turnovers of both OF system and O system. We observe that the two systems generate the same turnover at the end of the first day. At the end of the second day, the OF system is better than O system and at the end of the third day, the difference between the average turnovers of the two systems is even more important than the end of the second day. This observation confirms the results of the previous experiment, in the sense that, the difference between the average turnovers of the two systems evolves from "+20 euros" at the end of the second day to "+60 euros" at the end of the third day which confirms that the gain is more appreciated on a long period. In other word, the OF system has visibility on the Internet prospects' flows over the future periods that reports some sales in order to perform them more expensive later.
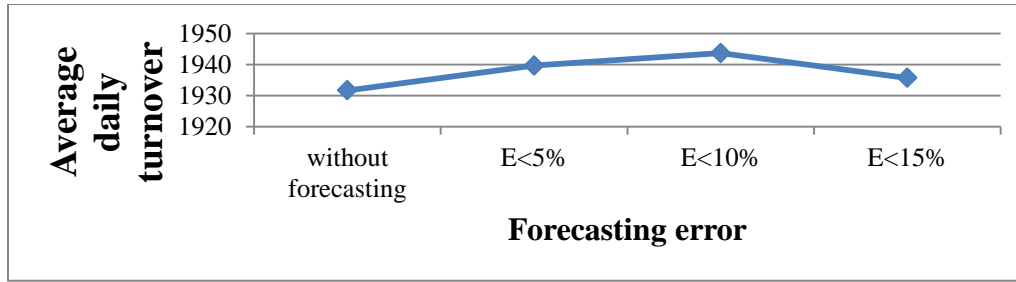
Figure 3: influence of the forecasting error rate on the average daily turnover

In the previous experiments, we studied the ability of the exact forecasting system to optimize the results of the multi-period assignment model. In this experiment, we address the influence of the forecasting error on the assignment model. Figure 3 shows that the integration of Internet prospects forecasting of the orders' groups having a forecasting error until 10% improves the results in terms of turnover. However, orders' groups with forecasting error more than 10% deteriorate the results.

5   CONCLUSION

In this paper, we presented a forecasting system that allows improving the results of a multi-period optimization model. Thus, the challenge behind this work was to determine the right data to forecast (not only a number of internet prospects) in order to integrate them into the optimization model. In other word, the proposed forecasting system returns the number of Internet prospects arriving at each period as well as their characteristics in order to determine orders for which these Internet prospects can be assigned. In addition, our forecasting system ensures the fact that forecasting's time step and optimization model's period must be the same in order to articulate both forecasting model and optimization model. Finally, in order to assess the benefit of using such an approach, empirical investigation was conducted on real data from PdL data base. Numerical tests proved that forecasting model is effective for realistic data amount. Hence, our contribution can be directly integrated into PdL platform as well as other real-world applications which are similar, for example, call centers and service companies.

6   REFERENCES

[1]   **Balinski, M.L.** A competitive (dual) simplex method for the assignment problem, Math. Program. 34 (2) (1986) 125–141.

[2]   **Box, G.E.P. Jenkins, G.M. Reinsel, G.C.** 2008. *Time Series Analysis: Forecasting and control*, 4th Edition, Wiley-Blackwell.

[3]   **Burkard, R. Dell'Amico, M. Martello, S. 2012**. Assignment Problems (revised version), *SIAM Society for Industrial and Applied Mathematics*, Philadelphia.

[4]   **M.M. Dessouky, B.A. Kijowski**, Production scheduling on single-stage multiproduct batch chemical process with fixed batch sized, IIE Trans. 29 (5) (1997) 399–408.

[5]   **Kuhn, H.W. 1955.** The Hungarian method for the assignment problem, *Naval Research Logistics Quarterly,* 2(1-2), 83-97.

[6]   **Lafaye De Micheaux, P. Drouilhet,R. Liquet, B.** 2014. *Le logiciel R,* 2nd Edition, Springer Verlag France.

[7]   **Maamar, M. Mousseau, V. Ouerdane, W. 2015.** Multi-period modeling and optimization of offer/demand assignment problem: a marketplace application, *unpublished manuscript*.

[8]   **F. Soumis, J. Ferland, J. Rousseau,** A model for large-scale aircraft routing and scheduling problems, Transport. Res. Part B: Meth. 14 (1) (1980) 191–201.