Proceedings of the 2015 ORSSA Annual Conference
pp. 23–32
www.orssa.org.za/wiki/uploads/Conf/2015ORSSAConferenceProceedings.pdf

ORSSA Proceedings
ISBN
978-1-86822-666-5


# Decision support for the assignment of real-estate agents to suburbs

J Janse van Rensburg*† JH van Vuuren‡

**Abstract**

Matching a real-estate agent to a suburb that best fits his or her property marketing expertise is an important (and difficult) decision for a real-estate agency aiming to achieve good turnover. In this respect the agents have to take into consideration the expected income that can be associated with a suburb when searching for new property stock to include in their personal portfolios. Real-estate agencies are also interested in having a presence in as many suburbs of a city as possible so as to be able to attract potential sellers in a large area. A bi-attribute decision support framework is proposed in this paper to aid real-estate agencies in selecting the best suburbs in such a way that both the agency's exposure and its expected income is maximised within a specified city. For this purpose a rich data set is required, containing all listed properties within the city, the agency's corresponding listing agents, as well as the values, locations and characteristics of properties in the market within the city. The working of the decision support framework is illustrated in the context of a special case study involving a real-estate agency in the Somerset West basin of the City of Cape Town, which employs eleven estate agents.

**Key words:** Real-estate, Agent assignment, Multiple criteria decision support, Suburb selection.

## 1 Introduction

Decision-making techniques can be separated into two broad categories: *group decision-making techniques* and *individual decision-making techniques*. There is a growing need by both individuals and businesses in a variety of sectors to use software in aid of good decision-making. This is due to increasing complexity associated with decision making as technology develops and data become abundantly available, thereby increasing the need to consider larger sets of stakeholders, criteria, alternatives and other factors that affect decisions.

*Department of Logistics, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, email: `johanjvrens@sunore.co.za`

†Stellenbosch Unit for Operations Research and Engineering, Department of Industrial Engineering, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa

‡**(Fellow of the Operations Research Society of South Africa)**, Stellenbosch Unit for Operations Research and Engineering, Department of Industrial Engineering, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, email: `vuuren@sun.ac.za`

The objective in this paper is to put forward a decision support framework for the assignment of real-estate agents to suburbs based on the available properties for sale as a function of time within the Somerset West basin. The system is aimed at the facilitation of agency-based decision-making in cases where the sets of assignment alternatives are large. The paper is a report on work in progress within a larger and ongoing research project at Stellenbosch University and is structured as follows. After performing a brief review of relevant literature in §2 on assignment problems and decision-making in the context mentioned above, we describe in §3 the problem considered in this paper and we turn our attention in §4 to a discussion on how the required case study real-estate data were captured and cleaned. This is followed, in §5, by a detailed proposal of how the algorithms associated with the multi-criteria assignment decision may be utilised to facilitate decision-making in the context of real-estate agent assignment to suburbs. An overview of the results obtained when applying the methodology to the data described in §4 is presented in §6. The paper closes in §7 with a brief conclusion and some pointers to related further work are given in §8.

## 2 Related literature

The assignment problem is one of the fundamental combinatorial optimisation problems in the operations research literature. It consists of finding a maximum weight matching (or minimum weight perfect matching) in a weighted bipartite graph. In its most general form, the problem may be described as follows. There are a number of agents and a number of tasks. Any agent can be assigned to perform any task, incurring some cost that may vary depending on the agent-task assignment. It is required to perform all tasks by assigning exactly one agent to each task in such a way that the total cost of the assignment is minimised.

If the numbers of agents and tasks are equal and the total cost of the assignment for all tasks is equal to the sum of the costs for each agent (or the sum of the costs for each task), then the problem is called the classical assignment problem [1]. A variety of algorithms have been devised to solve this assignment problem, such as the Hungarian algorithm [2] and the Auction algorithm [3]. There are even algorithms for solving multi-objective versions of the assignment problem [4].

In economics, diminishing returns (also called law-of-diminishing returns or marginal returns) is the decrease in the marginal (incremental) output of a production process as the amount of a single factor of production is incrementally increased, while the amounts of all other factors of production remain constant [6]. This simply means higher investment does not always fetch better returns. The law of diminishing returns states that as one invests more money, effort or time in an investment, the marginal rate of return eventually drops.

# 3 Problem description

A basic problem faced by real-estate agencies, and the problem considered in this paper, involves the selection of some subset of suburbs in which to invest the time, energy and resources of the agents employed by an agency. This problem is not only limited to the current staff of an agency, as agencies are typically also interested in the potential gains that may realise as a result of expanding their workforce. Finding the most valuable set of suburbs to focus on within a city is based on two criteria. The first criterion is to find a suburb set that maximises the agency's expected income, while the second criterion is to ensure that the agency is exposed to as many suburbs and subsequent properties which fall under its jurisdiction as possible in order to minimise the risk of its properties being withdrawn from the listings.

The following assumptions are made in this paper:

1. *All agents are equally likely to make a sale in any suburb.* This assumption is necessary in order to consider all suburbs as fertile ground for all the agents. It may, however, be relaxed by recording viable suburb sets per agent, based on their respective skill sets, and then limiting the suburb set assignable to each agent.

2. *All agents have similar workloads and are willing to work in any suburb.* This assumption negates the possibility that some agents may be assigned to suburbs containing only a few properties while other agents may be assigned to suburbs with a large collection of properties, in which case their workloads and respective incomes may differ vastly. The impact of this assumption can easily be reduced by assigning more than one agent to a suburb if there are too many properties in a suburb to be handled by a single agent, but this possibility lies beyond the scope of the current paper.

3. *The only agents competing in a specific suburb are the agents with listings in the suburb.* This assumption is made to simplify the problem under consideration, as the total number of agents is in reality typically unknown, but can be estimated by recording all suburbs' sales records per agent in the city so as to construct a suburb list per agent.

4. *All agents within the agency are to be assigned to one suburb only, but many agents may be assigned to any one suburb.*

In the case study considered in this paper, which involves the Somerset West basin (partitioned into 63 suburbs and comprising approximately 19 000 properties), only residential properties are considered. No *commercial properties*, *farms*, *apartments* or *plots* are therefore considered, as the assignment problem for the set of combined property types can be partitioned into smaller disjoint problems specific to each property type.

Of the 44 real-estate agencies operating in the Somerset West basin, one specific real-estate agency is considered as case study agency throughout this paper. This real-estate agency was selected as it already tracks all the information and data necessary to solve the problem at hand.

The problem considered here is different from the classical assignment problem, and even its multi-objective variants, in that the suburb properties remain the same for all agent assignments and that the agents and suburbs are not equal in number.

## 4 Data capturing and cleaning

Somerset West originally consisted of three separate areas, namely Bakershoogte, Parel Vallei and Somerset West, but today they are all merged to form the Somerset West basin. Due to this legacy divide, care must be taken when attempting to identify properties uniquely when working with their addresses, since an erf number may potentially refer to three different properties.

In what follows, a *property snapshot* refers to a *Microsoft Excel* file containing records of all properties that are in the market by all agencies in Somerset West. Such a data file is produced bi-weekly. The property snapshot data used in this paper were taken on 16 March 2015 and contained 571 properties, with 19 attributes tracked for each of these properties. For the purpose of this paper, only four of these attributes were used. They are the *advertised price* of the property, the *property address* (which includes the suburb in which it resides), the *agent* who has a mandate on the property and the *agency* for which the agent works.

The data set first had to be cleaned by removing all unusual characters and spacings. Critical fields were checked for completeness and cross-checked for accuracy and spelling. All the suburb field data were validated against *Google*'s map *application programming interface* (API). The agent details were cleaned and unique agency data were extracted corresponding to the case study agency mentioned above.

The programming language R was used to read the data files, clean the data and plot the graph included in this paper.

## 5 Modelling approach

Let $m$ denote the number of agents to be assigned to $n$ suburbs, and let

$$\boldsymbol{a}^{(i)} = \left[ \begin{array}{ccccc} a_1^{(i)} & \cdots & a_j^{(i)} & \cdots & a_n^{(i)} \end{array} \right], \quad i = 1, \ldots, \ell$$

denote the $i^{th}$ assignment alternative, where $a_j^{(i)}$ represents the number of agents assigned to suburb $j$ in alternative $i$.

The $i^{th}$ assignment alternative has the property that

$$\sum_{j=1}^{n} a_j^{(i)} = m, \quad i = 1, \ldots, \ell. \tag{1}$$

An upper bound

$$a_j^{(i)} \leq u, \quad i = 1, \ldots, \ell, \quad j = 1, \ldots, n. \tag{2}$$

is placed on the number of agents that may be assigned to any one suburb. The average agent income $x_j$ generated by suburb $j$ is the total property sales value $s_j$ in suburb $j$ divided by the total number of agents $\alpha_j$ (both internal and external to the agency) working in suburb $j$.

Let $y_j$ denote the total number of properties listed in suburb $j$ and let $r$ denote the agency's commission rate. The evaluation criteria associated with the $i^{th}$ assignment alternative may be expressed mathematically as

$$c_1^{(i)} = r\sum_{j=1}^{n} x_j a_j^{(i)}, \qquad i = 1, \dots, \ell, \tag{3}$$

and

$$c_2^{(i)} = \sum_{j=1}^{n} y_j a_j^{(i)}, \qquad i = 1, \dots, \ell. \tag{4}$$

The evaluation matrix

$$\boldsymbol{C} = \begin{bmatrix} c_1^{(1)} & c_2^{(1)} \\ \vdots & \vdots \\ c_1^{(i)} & c_2^{(i)} \\ \vdots & \vdots \\ c_1^{(\ell)} & c_2^{(\ell)} \end{bmatrix}$$

contains all quality evaluations associated with assignment alternatives with respect to the criteria in (3)–(4) and for all $i = 1, \dots, \ell$.

The expected agency income $E_j^{(i)}$ for the $i^{th}$ assignment alternative is measured as the average income $x_j$ per agent, weighted by the number of agency agents $a_j^{(i)}$ assigned to suburb $j$, *i.e.*

$$E_j^{(i)} = a_j^{(i)} \times x_j^{(i)}, \quad i = 1, \dots, \ell, \quad j = 1, \dots, n. \tag{5}$$

In order to forecast the expected return associated with assigning an additional agent to a suburb, the expected return may be calculated as the marginal return the newly assigned agent will add to the expected return for suburb $j$, taking into account the smaller share the already assigned agents will receive after the assignment.

As a result of the problem dimensions, we designed a heuristic for finding approximate solutions to the case study agent-suburb assignment problem instance. Given a sufficiently large second criterion value, our algorithm selects the best possible alternative $\boldsymbol{a}^{(i)}$ based on the first criterion which satisfies the lower bound threshold on the second criterion. This is achieved by sorting the expected agency income values $E_1^{(i)}, \dots, E_n^{(i)}$ in decreasing order, selecting the $m$ best possible assignments according to the first criterion and then testing whether the lower bound condition set on the second criterion is met by the assignment. The agent assignment is altered iteratively by replacing the worst assignment with respect to the first criterion with the best assignment for the second criterion that was not in the list, followed by the second best assignment, then the third best assignment, *etc.* until the second criterion's threshold value is reached.

Algorithm 1 may be used to generate an alternative assignment that will meet the lower bound selected for the second criterion.

**Algorithm 1** Generate an alternative assignment

```
sort data desending on criteria one's values
t ← select the lower bound for the second criterion
a ← first criterion data
b ← second criterion data
x ← assign the top m records of the sorted data
y ← assign the remaining n − m records
i = 0
for i in 1:nrow(x)-1 do
    if sum(x.b) >= t then break
    end if
    indexY ← which(yb == max(yb))[1]        ▷ select second best value criterion
    indexX ← nrow(x) − i
    xStor ← x[indexX, ]
    x ← rbind(x[−indexX, ], y[indexY, ])
    y ← rbind(y[−indexY, ], xStor)
end for
```

In cases where more than one agent is assigned to a suburb, the suburb may appear multiple times in the assignment matrix.

# 6 Case study results

The results presented in this section all refer to the 16 March 2015 data snapshot of listed properties for the case study agency, as described in §4.

The case study real-estate agency has the following attributes. Currently there are eleven active agents specialising in *residential property* at the agency. Together they currently cover 15 unique suburbs of the total of 63 suburbs; two suburbs are covered twice. These parameters lead to $\ell = \binom{n+m-1}{m} \approx 3.55 \times 10^{12}$ possible assignment alternatives in total. Some agents also focus on more than one suburb, resulting in a total of 17 possible suburb assignments. The current agency configuration of 17 assignments results in an expected income of R 3 139 126 and an exposure to 137 properties across 15 suburbs. These suburbs are shown in Table 1.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Audas | 3 | Bayview Heights | 7 | Briza | 11 | Die Wingerd |
| 14 | Fernwood Estate | 18 | Haumannshof | 23 | Heritage Park | 29 | La Sandra |
| 31 | Links | 39 | Raithby | 43 | Schonenberg | 47 | Somerset Mall |
| 49 | Somerset Ridge | 53 | Stuarts Hill | 60 | Winery Road | | |

**Table 1:** *The case study real-estate agency's 15 suburbs used for their 17 agent assignments, denoted in the text by* $\boldsymbol{a}^{(0)}$*, as it was on 16 March 2015.*

In the results reported in this section only eleven agent assignments are made in contrast to the case study agency's 17 agent assignments shown in Table 1. In other words, only one suburb is assigned per agent, while multiple agents may be assigned to a suburb (therefore possibly resulting in fewer than $m$ uniquely covered suburbs).

With only eleven agents, the largest possible value for the first criterion is obtained by the assignment of agents to the suburbs shown in Table 2, which achieves an expected income of R 9 289 561 and an exposure to 91 properties. This assignment was calculated by selecting the 11 best suburb agent assignments based on expected returns.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 12 | Erinvale Estate | 21 | Heldervue | 43 | Schonenberg | 50 | Spanish Farm |
| 51 | Stellenbosch | 61 | Worlds View | | | | |

**Table 2:** *The assignment alternative $\boldsymbol{a}^{(1)}$ of agents to the eleven suburbs yielding the largest possible value for the first assignment criterion.*

To calculate the upper bound for the second criterion the list of suburbs was sorted descending according to number of properties and the eleven highest suburbs was selected. These suburbs are shown in Table 3. The best assignment of agents to these eleven suburbs achieves an exposure to 269 properties and an expected return of R 6 666 329.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Audas | 4 | Bizweni | 12 | Erinvale Estate | 21 | Heldervue |
| 22 | Helena Heights | 28 | La Concorde | 33 | Montclair | 37 | Parel Vallei |
| 43 | Schonenberg | 44 | Sir Lowry's Pass | 50 | Spanish Farm | | |

**Table 3:** *The assignment $\boldsymbol{a}^{(2)}$ of agents to the eleven suburbs yielding the largest possible value for the second assignment criterion.*

The recommended assignment of agents to the suburbs in Table 4 was calculated by replacing the lowest expected return in the list of assignments in $\boldsymbol{a}^{(1)}$ with the second, third, *etc.* lowest expected return until the desired value for criterion 1 had been reached. For the case study agency the best assignment of agents to these properties resulted in an exposure to 151 properties and an expected return of R 9 197 771.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 12 | Erinvale Estate | 21 | Heldervue | 37 | Parel Vallei | 43 | Schonenberg |
| 50 | Spanish Farm | 51 | Stellenbosch | 61 | Worlds View | | |

**Table 4:** *The assignment alternative $\boldsymbol{a}^{(3)}$ achieving the best possible value for the second assignment criterion with a first assignment criterion larger than or equal to that of $\boldsymbol{a}^{(0)}$.*

The assignment alternative for the current agency suburb selection is

$$\boldsymbol{a}^{(0)} = \left[\begin{array}{ccccccc} a_1^{(0)} = 2 & a_3^{(0)} = 1 & a_7^{(0)} = 1 & a_{11}^{(0)} = 1 & a_{14}^{(0)} = 1 & a_{18}^{(0)} = 1 & a_{23}^{(0)} = 1 \\ a_{29}^{(0)} = 1 & a_{31}^{(0)} = 1 & a_{39}^{(0)} = 1 & a_{47}^{(0)} = 1 & a_{49}^{(0)} = 1 & a_{53}^{(0)} = 1 & a_{60}^{(0)} = 2 \end{array}\right]$$

with corresponding evaluation matrix entries

$$\boldsymbol{C}_0 = \left[\begin{array}{cc} c_1^{(0)} = 3\,139\,126 & c_2^{(0)} = 137 \end{array}\right],$$

where we follow the convention that all $a$-values not listed assume the value zero. The assignment alternative achieving the largest income was, however, found to be

$$\boldsymbol{a}^{(1)} = \left[\begin{array}{cccccc} a_{12}^{(1)} = 5 & a_{21}^{(1)} = 2 & a_{43}^{(1)} = 1 & a_{50}^{(1)} = 1 & a_{51}^{(1)} = 1 & a_{61}^{(1)} = 1 \end{array}\right]$$

with corresponding evaluation matrix entries

$$\boldsymbol{C}_1 = \left[\begin{array}{cc} c_1^{(1)} = 9\,289\,561 & c_2^{(1)} = 91 \end{array}\right],$$

while the assignment alternative achieving the largest exposure is

$$\boldsymbol{a}^{(2)} = \left[\begin{array}{cccccc} a_1^{(2)} = 1 & a_4^{(2)} = 1 & a_{12}^{(2)} = 1 & a_{21}^{(2)} = 1 & a_{22}^{(2)} = 1 & a_{28}^{(2)} = 1 \\ a_{33}^{(2)} = 1 & a_{37}^{(2)} = 1 & a_{43}^{(2)} = 1 & a_{44}^{(2)} = 1 & a_{50}^{(2)} = 1 & \end{array}\right]$$

with corresponding evaluation matrix entries

$$\boldsymbol{C}_2 = \left[\begin{array}{cc} c_1^{(2)} = 6\,666\,329 & c_2^{(2)} = 269 \end{array}\right].$$

The agency was quite happy with their current exposure and sought the best possible expected return with a similar or better exposure. Hence, we recommended the assignment

$$\boldsymbol{a}^{(3)} = \left[\begin{array}{ccccccc} a_{12}^{(3)} = 4 & a_{21}^{(3)} = 1 & a_{37}^{(3)} = 1 & a_{43}^{(3)} = 1 & a_{50}^{(3)} = 2 & a_{51}^{(3)} = 1 & a_{61}^{(3)} = 1 \end{array}\right]$$

with corresponding evaluation matrix entries

$$\boldsymbol{C}_3 = \left[\begin{array}{cc} c_1^{(3)} = 9\,197\,771 & c_2^{(3)} = 151 \end{array}\right],$$

which yields an expected return of 193% of the current expected return with a 10% increase in property exposure.

The results described above are summarised in Table 5 and illustrated graphically in Figure 1.

| $\boldsymbol{a}^{(i)}$ | Total suburb expected income | Total property exposure | Description | Status |
|---|---|---|---|---|
| $\boldsymbol{a}^{(0)}$ | R 3 139 126 | 137 | current | dominated |
| $\boldsymbol{a}^{(1)}$ | R 9 289 561 | 91 | highest earnings | non-dominated |
| $\boldsymbol{a}^{(2)}$ | R 6 666 329 | 269 | highest exposure | non-dominated |
| $\boldsymbol{a}^{(3)}$ | R 9 197 771 | 151 | recommended | non-dominated |

**Table 5:** *Relative performances of four assignment alternatives.*

## 7 Conclusion

The president of the case study agency's intuition going into this research endeavour was that 12 Erinvale Estate, 37 Parel Vallei, 43 Schonenberg and 50 Spanish Farm are the suburbs that are worth spending extra time and money on.

The results obtained in §6 were presented to the president and the recommended alternative $\boldsymbol{a}^{(3)}$ aligned with what the agency thought might be the most profitable suburbs. It is interesting to note how similar the recommended assignment is to the intuition of the agency's president. Although the assumptions made in §3 seem to be rather restrictive, the results are quite encouraging.

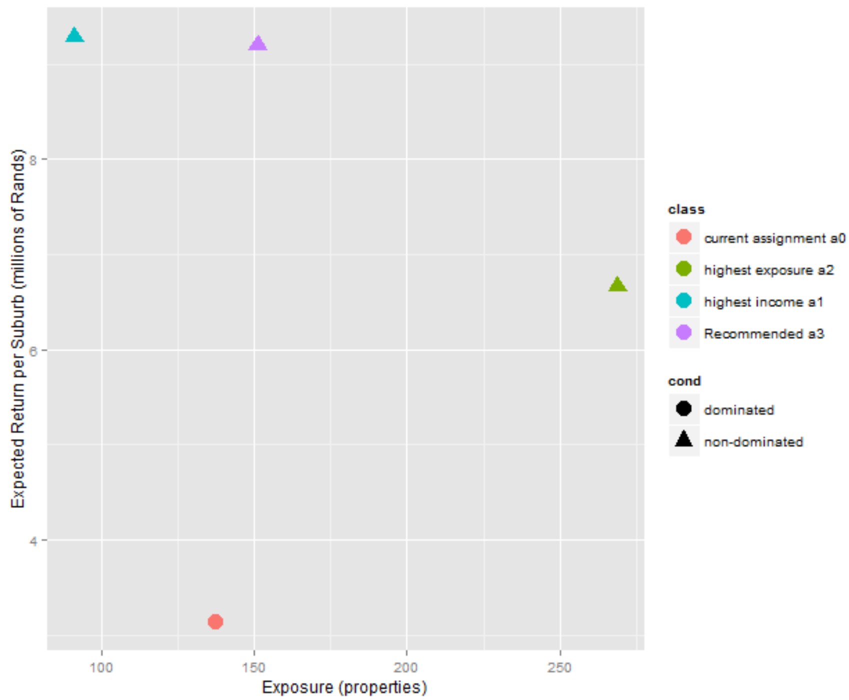


**Figure 1:** *The real-estate agency's current assignment is suboptimal compared to the approximately Pareto set, shown as triangles. Both the best possible expected return in (3) and the best possible exposure in (4), are shown.*

## 8 Further work

This paper only touched on the basics of the fundamental problem of assigning agents to suburbs. Much more can, however, be done to refine the modelling approach so as to better reflect real-world exposure and expected return (such as, for example, taking into account the expertise and marketing skills of agents). The results obtained in §6 may also be compared with those returned by other (meta)heuristic and MCDA approaches, such as, for instance, selecting an approximately Pareto-optimal set of solutions generated by an evolutionary algorithm. Historical data may further be used to predict popular suburbs for future assignments.

## References


[1] Munkres J, 1957, *Algorithms for the assignment and transportation problems*, Journal of the Society for Industrial and Applied Mathematics, **5(1)**, pp. 32–38.

[2] Kuhn HW, 1991, *On the origin of the Hungarian Method, history of mathematical programming*, North Holland, Amsterdam, pp. 77–81.

[3] Bertsekas DP, 1988, *The auction algorithm: A distributed relaxation method for the assignment problem*, Annals of Operations Research, **14(1)**, pp. 105–123.

[4] De PK & Yadav B, 2011, *An algorithm to solve multi-objective assignment problem using interactive fuzzy goal programming approach*, International Journal of Contemporary Mathematical Sciences, **6(34)**, pp. 1651–1662

[5] Weistroffer HR & Narula SC, 2005, *Multiple criteria decision support software*, Springer, New York (NY).

[6] Shephard RW, 1969, *Proof of the law of diminishing returns*, University of California, Berkeley.