Matching models for evacuation and allocation of people in case of disasters and wars

Written by Simone Bastasin[1,2], Simone Bortolin[1,3], Thomas Trevisan[1,4]

[1]Department of Information Engineering - University of Padua, Italy

[2]simone.bastasin@studenti.unipd.it, [3]simone.bortolin.1@studenti.unipt.it, [4]thomas.trevisan@studenti.unipd.it

Abstract

In this article we analyse an innovative case study in machine learning: matching models for evacuation and allocation of people in case of disasters and wars.

We will briefly introduce matching under preferences and in particular *matching model with sizes* because it fits very well for our case study. After defining both models we will analyse which algorithms perform better to solve the two matching problems.

Introduction

The topics covered in this paper have been chosen mainly for two nowadays problems. First, as described by Olberg and Seuken (2022), the Swiss State Secretariat for Migration recently announced an innovative machine learning-based assignment process for refugee resettlement that considers also refugees' preferences. Second, the increasing number of wars and terrorist attacks have highlighted the necessity to intelligently manage the evacuation of people from dangerous places in a more effective way.

Moreover, in the last period, due to the increase of natural catastrophes and to the war in Ukraine it becomes even more necessary to have matching models that manage automatically these kinds of problems. At first sight they might seem unrelated each other but in reality they are extremely tied and similar.

Matching model under preferences

Matching under preferences is a tool from cooperative game theory. It can be applied to two-sided markets in which heterogeneous agents (or objects) of one side are distributed over agents of the other side, considering the satisfaction of agents' preferences (or objects' priorities). Gale and Shapley (1962) formalized the theoretical foundations of this tool.

Matching markets with sizes arise in a variety of contexts, and in its simplest model, agents have ordinal preferences over objects that are available in multiple identical units. For each object, agents are ranked according to exogenous priorities. The model can be extended to fit specific applications if agents have preferences over both an object and a number of units.

An agent and an object form a *blocking pair* of a given matching if the agent prefers the object to his current one and the number of units he requires are assigned to agents with a lower priority. The number of units required has to be less or equal to the availability of the object.

Stability Stability, initially introduced by Gale and Shapley (1962), is a main concept in matching theory. A matching is *stable* if it doesn't have any blocking pair. In two-sided matching markets, with strategic agents on both sides, stability constitutes an essential equilibrium criterion. A blocking pair is unfair in the sense that the agent is not able to get an object even though he has a high enough priority.

Size envy-free A matching is said to be *size envy-free* if, whenever an agent prefers an object to his current one, all agents matched to that object have either a higher priority or a smaller size. Size envy-freeness is a fairness criterion in the sense that any priority violation can be justified by the different sizes of the agents. *Weak envy-freeness* relaxes envy-freeness by allowing some innocuous priority violations.

Non-wasteful A matching is defined *non-wasteful* if, whenever an agent prefers an object to his current one, the object has enough unassigned units to be matched to the agent without removing any other agent. Non-wastefulness constitutes both a fairness and an efficiency criterion as it ensures that units only remain unassigned if they cannot benefit any agent.

Logical relationships Weak envy-freeness is logically independent of stability in the same way as envy-freeness is. In contrast, stability only fails due to the presence of waste or via a priority violation so if a matching is non-wasteful and weakly envy-free, then it's stable. A graphic representation of the relationships between solutions is present in the Figure 1.

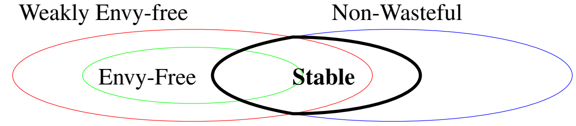

Figure 1: Logical relationships.

Matching models

Matching model for refugee resettlement

To avoid that refugee families seeking shelter are assigned to countries randomly, Bansak et al. (2018) propose a machine learning-based algorithm for family placement that aims to optimize the overall employment rate of refugees. However, this method ignores families' preferences.

Olberg and Seuken (2022) propose two matching mechanisms that, additionally to the optimization of the employment success, take into account also refugees' preferences over locations. These innovative methods improves average resettlement success, allowing a trade off between family welfare and overall employment success.

According also to the formally notation of Salles (2017) and Delacrétaz, Kominers, and Teytelboym (2020), an instance of a refugees-countries matching problem is a 6-tuple (C, R, q, P_c, P_r, F), where $C = \{c_1, \ldots, c_m\}$ and $R = \{r_1, \ldots, r_n\}$ are disjoint sets of m countries and n refugees, respectively.

Since the distribution for refugee resettlement is modelled as a matching problem under preferences, refugees and countries offering asylum can be considered as a two-sided market in which the members of one side are distributed over the members of the other side. So, we can define the agents of the market as $a_k \in C \cup R$. Note that we are concerned with *many-to-one matchings* since it can be assumed that $n \gg m$ and each refugee can obtain asylum in at most one country, whereas a given country can accept many refugees.

The maximum number of refugees that can be matched to each country is determined by a vector of quotas $q = (q_j)_j \in \mathbb{N}^m,\ j \in \{1, \ldots, m\}$. There may be no real quotas at all: setting $q_j = n\ \forall c_j \in C$ makes them dummies.

Afterwards, $P_c = \{P(c_1), \ldots, P(c_m)\}$ and $P_r = \{P(r_1), \ldots, P(r_n)\}$ are sets of preference lists which include a complete and transitive preference profile for each country over the set of refugees and for each refugee over the set of countries. Each preference $P_r(r_i)$ contains a list of expressed preferences in the format $c_1 \succ_{r_i} c_2$, and equivalently for countries' preferences. An example is in Table 1.

In the particular case of refugees' allocation, last elements to consider are the groups of people who must be assigned together (e.g. a family): let $F = \{F(f_1), \ldots, F(f_l)\}$ be the list of the groups of refugees, where $F(f_i) = \{r_a, r_b, \ldots\}$ and $l \leq n$. An example can be found in Figure 2.

A country may declare some refugees unacceptable and a refugee may declare some countries unacceptable, hence, $E \subseteq R \times C \times F$ is the subset of the acceptable refugee-country pairs. In addition, the amount of acceptable tuples is bounded because, as a refugee can't be split into two nations, nor can one nation receive more than the maximum allowable quotas. Denote $A(r_i) = \{c_j \mid (r_i, c_j) \in E\}$ as the set of acceptable countries for a given $r_i \in R$; and equivalently for the countries.

An assignment M is a subset of E that contains $a_k \in R \cup C$ items. Obviously a refugee r_i can be unassigned: $M(r_i) = \emptyset$. Similarly, a country c_j can admit asylum requests if $|M(c_j)| < q_j$ and, therefore, requests are blocked if $|M(c_j)| = q_j$. Note that the assignment is valid if and only if $|M(r_i)| \leq 1\ \forall r_i \in R$ and $|M(c_j)| \leq q_j\ \forall c_j \in C$.

Countries	Refugees
$r_1 \succ_{c_1} r_2 \succ_{c_1} r_3$	$c_2 \succ c_1$ for both r_1 and r_2
$r_2 \succ_{c_2} r_1 \succ_{c_2} r_3$	r_3 declares only c_2 acceptable

Table 1: Refugees and countries' preferences example. Refugees r_1, r_2, r_3; countries c_1, c_2 with $q_1 = 2,\ q_2 = 1$. Notation $a \succ_c b$ denotes that c strictly prefers a to b.

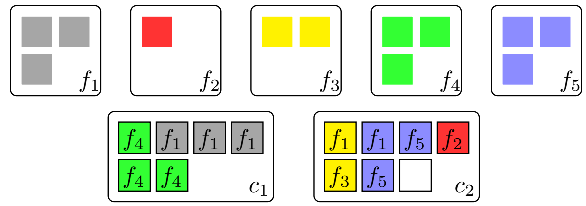

Figure 2: Full matching example. Families $f_1, \ldots, f_5$ with $|f_1| = |f_4| = |f_5| = 3,\ |f_2| = 1,\ |f_3| = 2$; countries c_1, c_2 with $q_1 = 6,\ q_2 = 7$.

Matching model in case of disaster

In the event of a catastrophe (fire, earthquake, hurricane, bomb, ...) it's necessary to study and implement an evacuation system capable to guarantee an efficient and orderly escape, causing the least panic possible. In particular, an evacuation model can be identified by two elements: people and escape routes.

As established by Italian Decreto Legislativo n. 81/2008, American Title 29 of the Code of Federal Regulations (CFR) 1910.33-39 and European Directive 89/654/EEC, each state has its own regulations that can be summarised as shown in the Figure 3.

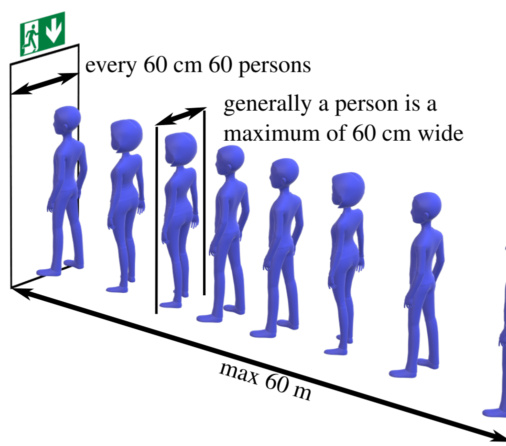

Figure 3: Principles underlying the evacuation model: every 60 cm width of a door, corridor or staircase[1] can come out around 60 people[2] and from the middle of a room to the nearest emergency exit there can be maximum 60 meters.

[1]60 cm is the minimum width of a passageway for a person to walk easily without crawling or colliding.

[2]Regard regulations, in America this number is always fixed at 60, in Europe is more variable, usually it's 50 but it can go up to 70 in education facilities or private offices. For example, a 120 cm wide escape route can allow the exit of 100–140 people, and a 90 cm escape route can allow the exit of 50–70 people.

The problem of finding the best escape route for every person can be viewed as a matching problem very similar to the one for refugee resettlement[3]. An instance of a people-exit matching problem is a 6-tuple (E, P, q, N_e, N_p, F), where $E = \{e_1, \dots, e_m\}$ and $P = \{p_1, \dots, p_n\}$ are disjoint sets of m exits and n people, respectively.

For the same reasons as before, we can define the agents of the market as $a_k \in P \subset E$ and, since it can be assumed that $n \gg m$, we are concerned with many-to-one matchings.

The maximum number of people that can be matched to each exit is determined by a vector of quotas $q = (q_j)_j \in \mathbb{N}^m$, $j \in 1, \dots, m$.

Afterwards, $N_e = \{N(e_1), \dots, N(e_m)\}$ and $N_p = \{N(p_1), \dots, N(p_n)\}$ are sets of preference lists which include a complete and transitive preference profile for each exit over the set of people and for each person over the set of exits. Each $N_p(p_i)$ contains a list of exits ordered in the format $e_1 \succ_{p_i} e_2$ where e_1 is nearer than e_2. Similarly, each $N_e(e_i)$ contains a list of people ordered by priority (e.g. a staircase with ramp gives priority to people with physical disabilities). An example is illustrated in Table 2.

Similarly to the previous model, last elements to consider are the groups of people who must escape together (e.g. a classroom): let $F = \{F(f_1), \dots, F(f_l)\}$ be the list of these groups, where $F(f_i) = \{p_a, p_b, \dots\}$ and $l \leq n$.

The subset of acceptable person-exit pairs is defined as $C \subseteq E \times P \times F$. In addition, commonly to the previous model, the amount of acceptable tuples is bounded because, as a person can't be split into two exits, nor can one exit receive more than the maximum allowable quotas. Denote $A(e_i) = \{p_j \mid (e_i, p_j) \in C\}$ as the set of acceptable people for a given $e_i \in E$; and equivalently for the people.

An assignment M is a subset of C and contains the item $a_k \in E \cup P$. Obviously in case of a disaster every person should be able to evacuate, so p_i can't be unassigned: $M(p_i) \neq \emptyset$.[4] Note that the assignment is valid if and only if: $|M(p_i)| = 1 \; \forall p_i \in P$ and $|M(e_j)| \leq q_j \; \forall e_j \in E$.

This model can help people a lot and, in particular, it can be widely used to manage evacuation and shelter to safe places in many locations and for various situations. For example, it can be applied to keep safe schools, hospitals, companies and campuses from dangers like earthquakes, fires, wars and severe floods.

People	Exit
$p_1 \succ_{e_1} p_2 \succ_{e_1} p_3$	$e_2 \succ e_1$ for both p_1 and p_2
$p_2 \succ_{e_2} p_1 \succ_{e_2} p_3$	p_3 declares only e_2 within 60 mt

Table 2: People and exits' preferences example. People, p_1, p_2, p_3, exits e_1, e_2 with $q_1 = 2,\ q_2 = 1$.

[3]The exit model can be used to make a simple escape plan in a static way using upper bounds on the affluence. Or it can be solved dynamically using real affluence data through a computer vision system and dynamic people-addressing systems.

[4]Although it's possible that for some reason an exit is impracticable and, just in this cases, it's possible to realize models in which you have $|M(p_i)| \geq 1$. In this way, to each user is assigned a second emergency exit, so that if the first one is impracticable the second one can be used. This strategy should be avoided and it's better to have a detection system for impracticable exits.

Use of matching models in case of war

Previous models argue topics that can be very useful in case of war. In particular, for mass escapes, first model can redistribute in a quick and fair way people to safe countries. Then, during bombardments or alarms of incoming attacks, second model can optimally manage:

- evacuation of people from buildings at risk of attack;
- shelter of people towards bunkers or other safe places.

Similarities and differences between mechanisms As described in this paper, the two proposed models, except for few differences, are very similar. For this reason it's possible to use a common system to analyse and solve the two matching problems. Delacrétaz, Kominers, and Teytelboym (2020) formulated a generalized model that can be adapted in many situations, like ours.

Mechanisms

Matching models as College Admissions Problem

A CA-instance of a refugees-countries matching problem is a 4-tuple (C, R, q, P) where $P = \{P_c, P_r\}$ (and F is omitted). Similarly, a CA-instance of a exits-people matching problem is a 4-tuple (E, P, q, N) where $N = \{N_e, N_p\}$.

We can now turn to the desiderata encountered in the previous section. It will be argued that the desideratum for which preferences are satisfied "as much as possible" amounts to stability in the CA model.

This model meets the requirements of stability and maximum cardinality, but for the case study of refugee resettlement, even if the requirement of "no rights violation" is met and it's adopted a modified CA model in which "everyone finds everyone acceptable", there is another problem to be taken into account: the system should ban the possibility of "incorrect use of preferences". An "incorrect use of preferences" is an expression of preferences which are in conflict with system compliance (COM): in this case higher-order policy goals or ethical principles. This happens because countries can express indirect preferences which can lead to discrimination (e.g. requiring language skills). As pointed out by (Van Basshuysen 2017), compliance is also violated if agents can "game the system".

Matching models as School Choice Problem

The SC model can be derived from the CA model by restricting the set of preference lists to only those of the refugees and defining priority lists for the countries. Thus, a SC instance of a refugees-countries matching problem is a 5-tuple (R, C, q, P_r, Pri) where $Pri = \{Pri(c_1), \dots, Pri(c_m)\}$ is the set of countries' priority lists containing same data of P_c.

Priority rankings are usually generated through a points system: if two applicants have identical points, the priority ranking may be determined through a lottery or continuous factors. In the context of refugee resettlement, to keep fairness between two refugees with same points, the refugee who has been waiting longer for asylum is prioritized. So, the waiting time since the asylum request could be used as a continuous variable to break non-strict priorities.

The SC model respects the requirements of stability, maximum cardinality and COM because it doesn't allow the inclusion of requirements contrary to higher-order policy goals or ethical principles (Van Basshuysen 2017).

Machine learning-based Matching models

Machine learning-based algorithms used for solving our case study can be viewed as:

- a clustering problem taking into account only people's preferences;
- a multilabel classification problem where every person is assigned to a country or an exit according to P_c or N_e.

Both approaches violate part of the preferences and COM but combining the result of the multilabel classification with some optimization algorithms it's possible to achieve interesting results. Let π_{ij} be the probability that the family i integrates properly in the country j, estimated from P_c using a machine learning process. We call the expected number of successfully integrated families, i.e., $z(\mu) = \sum_{i \in F} \pi_{i\mu(i)}$, the objective function of μ. A feasible matching μ^* is optimal if $\mu^* \in \arg\max_{\mu}\{z(\mu)\}$. Variable x_{ij} indicates whether the family i will be assigned to country j.

Our goal is to find a feasible matching μ that maximizes family welfare in terms of reported preferences P_r and at the same time ensures that $z(\mu)$ is within a factor of α of $z(\mu^*)$ for a previously chosen $\alpha \in [0,1]$.

Data: F, C, P_r, π, α
Result: μ
$\mu(i) \leftarrow \emptyset$ for all $i \in F$.

$$\begin{aligned} z^* \leftarrow \text{maximize} & \sum_{i \in F} \sum_{j \in C} \pi_{ij} a_{ij} \\ \text{subject to} & \sum_{i \in F} a_{ij} = q_j && \forall j \in C \\ & \sum_{j \in C} a_{ij} = 1 && a_{i\mu(i)} = 1 \\ & a_{i\mu(i)} = 1 && \forall i \in F(\mu) \\ & a_{ij} \in \{0,1\} && \forall i \in F, \forall j \in C \end{aligned}$$

$Q \leftarrow F$.
while $Q \neq \emptyset$ **do**
  Remove randomly chosen $i \in Q$ from Q.
  if $|F_j(\mu)| < q_j$ **then**
    $\mu' \leftarrow \mu; \mu'(i) \leftarrow j$

$$\begin{aligned} z' \leftarrow \text{maximize} & \sum_{i \in F} \sum_{j \in C} \pi_{ij} a_{ij} \\ \text{subject to} & \sum_{i \in F} a_{ij} = q_j && \forall j \in C \\ & \sum_{j \in C} a_{ij} = 1 && a_{i\mu(i)} = 1 \\ & a_{i\mu(i)} = 1 && \forall i \in F(\mu) \\ & a_{ij} \in \{0,1\} && \forall i \in F, \forall j \in C \end{aligned}$$

    if $z' \geq \alpha z^*$ **then**
      $\mu(i) \leftarrow j$
      break
    end
  end
end

Algorithm 1: Constrained Random Serial Dictatorship (CRSD)

A CRSD instance is $\hat{I} = (F, L, q, \pi, P, \gamma)$ and it returns $\mu : F \rightarrow C \cup \{\emptyset\}$. The general idea of CRSD is to let a family choose their preferences only from the set of remaining countries if it can be guaranteed an α-approximation of the objective. The constrained random serial dictatorship (CRSD) mechanism (Bansak et al. 2018) is a constrained version of the well known random serial dictatorship mechanism. Pseudocode of CRSD is illustrated in the Algorithm 1.

Constrained rank value (CRV) mechanism is a variant of the CRSD one which uses a rank function. It's based on $\hat{I} = (F, C, P_r, \pi, v, \alpha)$, where v is a rank function that assigns values between 0 and 1 to positions in preference order that is monotonically decreasing (Olberg and Seuken 2022). Pseudocode of CRV is illustrated in the Algorithm 2.

The machine learning model π is Admissible for COM because the algorithm used, even if it calculates a wrong value, doesn't cause any death.

Data: $F, C, P_r, \pi, v, \alpha$
Result: μ

$$\begin{aligned} z(\mu^*) \leftarrow \text{maximize} & \sum_{i \in F} \sum_{j \in C} \pi_{ij} a_{ij} \\ \text{subject to} & \sum_{i \in F} a_{ij} = q_j && \forall j \in C \\ & \sum_{j \in C} a_{ij} = 1 && a_{i\mu(i)} = 1 \\ & a_{i\mu(i)} = 1 && \forall i \in F(\mu) \\ & a_{ij} \in \{0,1\} && \forall i \in F, \forall j \in C \end{aligned}$$

$$\begin{aligned} \mu \leftarrow \text{maximize} & \sum_{i \in F} \sum_{j \in C} v\left(P_r(r_i)_j\right) a_{ij} \\ \text{subject to} & \sum_{i \in F} a_{ij} = q_j && \forall j \in C \\ & \sum_{j \in L} a_{ij} = 1 && \forall i \in F \\ & \sum_{i \in F} \sum_{j \in C} \pi_{ij} a_{ij} \geq \alpha z(\mu^*) \\ & a_{ij} \in \{0,1\} && \forall i \in F, \forall j \in C \end{aligned}$$

Algorithm 2: Constrained Rank Value Mechanism (CRV)

Conclusions

In this paper, we have analyzed how evacuation and allocation of people in case of disasters and wars can be dealt using artificial intelligence and, in particular, matching theory.

We have observed that, discarding models that don't respect ethic principles, for our case of study only two mechanisms are adaptable: School Choice Problem and Machine learning-based Matching. The latter model is an optimization model based on the contribution of machine learning and, since its output value π is very similar to the list of priorities of the School Choice Problem, we can say that in both models COM is not violated.

The main advantage of a matching model that can be adapted in various situations is that it allows you to work and improve a single model that can be reused also for many other purposes. In this way, it's possible to reduce effort on building more stable matching resolution algorithms.

Finally, we have seen how these models can be very useful if applied in the particular case of a global issue such as war that, unfortunately, is still very current. In fact:

- 1^{st} model can redistribute people in safe countries;
- 2^{nd} model can manage the evacuation from buildings at risk and the shelter to bunkers or other safe places.

References

Bansak, K.; Ferwerda, J.; Hainmueller, J.; Dillon, A.; Hangartner, D.; Lawrence, D.; and Weinstein, J. 2018. Improving refugee integration through data-driven algorithmic assignment. *Science* 359(6373): 325–329.

Council Directive of European Economic Community. 1989. Directive 89/654/EEC.

Delacrétaz, D.; Kominers, S. D.; and Teytelboym, A. 2020. Matching Mechanisms for Refugee Resettlement (2019-078).

Federal Government of United States of America. 1974. Exit Routes, Emergency Action Plans, and Fire Prevention Plans in Title 29 of the Code of Federal Regulations (CFR) 1910.33-39.

Gale, D.; and Shapley, L. S. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69(1): 9.

Olberg, N.; and Seuken, S. 2022. Enabling Trade-offs in Machine Learning-based Matching for Refugee Resettlement (4).

Prime Minister of the Italian Government. 2008. Decreto Legistrativo n. 81/2008.

Salles, M. 2017. Felix Brandt, Vincent Conitzer, Ulle Endriss, Jerôme Lang, and Ariel Procaccia (eds), Handbook of Computational Social Choice. *OEconomia* (7-4): 609–618.

Van Basshuysen, P. 2017. Towards a Fair Distribution Mechanism for Asylum. *Games* 8(4). ISSN 2073-4336.