

Recommending Franchises to the client on the basis of his available resources

SI583: Recommender Systems

SI, University of Michigan

Winter, 2010

Author: Divye Bokdia

ABSTRACT

This paper addresses the system design of a recommender system that will recommend Franchise to the Franchisees. There are about 3500 different Franchises available to choose from and all the Franchisors have their own expectations and requirements from the Franchises owner. As Franchisees have to totally rely on the knowledge of the Consultant and so there's lots of room for manipulate the Franchise recommended by them. This recommender system uses hybrid collaborative filtering algorithm to recommend list of Franchises. It matches both the interests of the Franchisees and their available resources, and what Franchises were opted by other clients with similar requirements. It also takes, error in recommendation, into account and updates the next recommendation accordingly.

Introduction

There are many consultancy firms, which help their clients to decide and own a Franchise based on their resources like time they can give, office space they have, money they can invest, number of employees they can hire and afford, and the most important, business domain they are interested in like medical, health care, food, automobile et cetera. The consultants at such firms are experienced and have a broad knowledge about each of the Franchisors and Franchise.

Definitions¹

Following are some of the terms that we will be using frequently.

Franchisor: One that grants a Franchise.

Franchise: Authorization granted to someone to sell or distribute a company's goods or services in a certain area.

Franchisee: One that is granted a Franchise, as to market a company's goods or services in a certain local area. (*Our client*)

Business Constrains

Firstly, Currently Consultants assess the requirements of the client and then recommends the Franchise to them, based on their experience, understanding and knowledge. Thus, the recommendations are totally based on their knowledge and they have to select among 3500 Franchises, so there are lots of chances that customer might not get what he wanted in the first shot itself.

Secondly, In Business, clients always have an option to shift from one firm to another. So if the Consultant is not able to deliver expected results, they might lose a customer. Thus consultants have to be very precise and flawless at the same time.

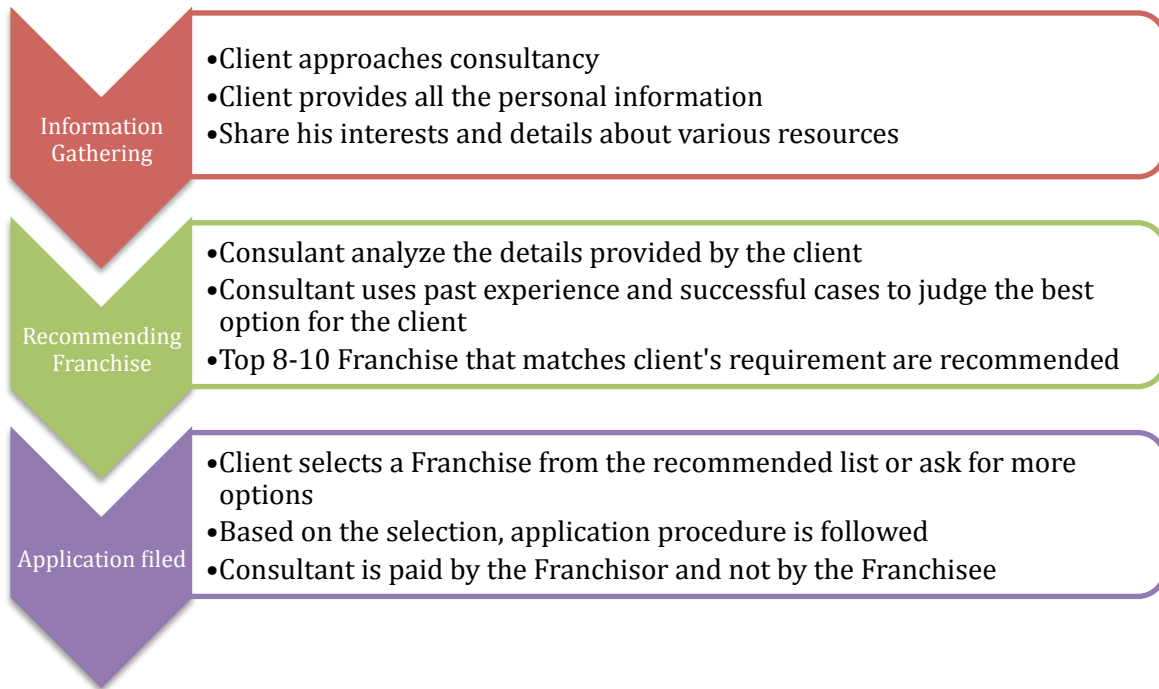
Thirdly, Any error made in the recommendations, can highly affect the business. As this is a matter of high monetary investment and thus a wrong suggestion from the consultant could lead to a big failure for the client. As the client who is interested in owning a Fitness Franchise, won't be satisfied and comfortable in owning a Coffee Shop Franchise.

Fourthly, The requirements for each Franchise are different like Coffee Franchise needs less investment and workers but demands prominent location. So all these factors are pivotal in deciding that which franchise will suits him the best.

Finally, Apart from the quantifiable details like investment available, office space, employees client can hire, time he can give etc there are some other aspects like the domains he is interested in or what are his interests. Now his interests can't be quantified and thus have to be considered very carefully.

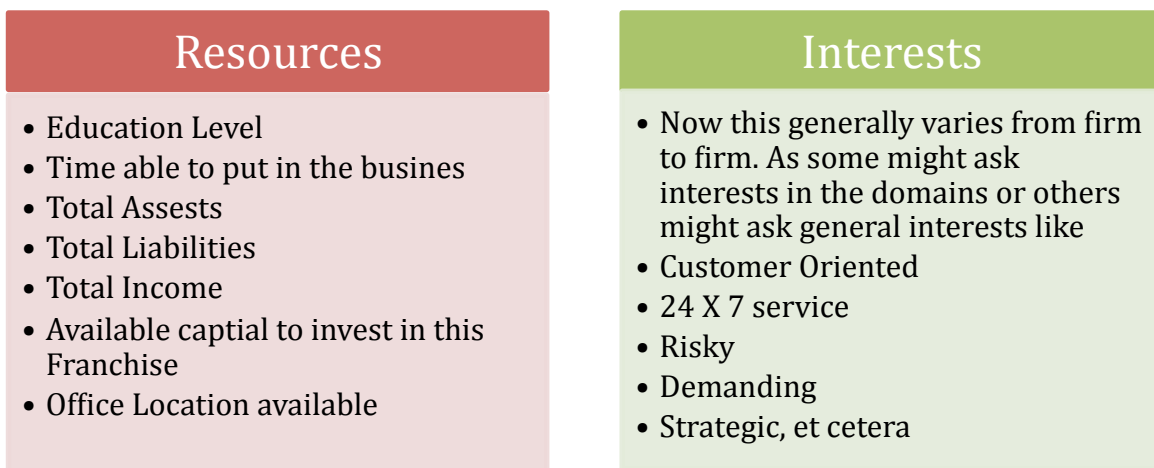
Franchise Consultancy

In a typical setup, following is the procedure followed between any client and a consultant.



Information from the Client

This is the most important phase of the process, as rest all the steps depends on what the client's needs are. Firms gather explicit information from the client under following two broad categories.



They gather all the information about the **available resources** and this is the first and primary criterion of Franchise selection. As if the Franchisee doesn't have resources to own a Franchise, then it doesn't make sense to recommend him that.

The other set of information gathered is client's **interests**. Now these interests could be from two different categories, first could be the type of domains he is interested in and second could be other abstract level interests like he wants something independent,

which involves talking to people and must be risky. The decision to collect information from which category (i.e. domain or general interest) is taken by the firm and they might even merge the interests from both the categories.

Recommender Algorithm

Two-tier Hybrid Algorithm

This is an interesting situation and no single algorithm will give us the most appropriate results here. We can't apply only User-User algorithm to find all the users who have similar values for the resources and then recommend the same list of Franchises (that were recommended to similar user) to the new user. They might not have the same level of matching in values for the resources and interests at the same time.

Lets say each user has to provide value for all the 7 resources (R1, R2, ...R7) and he can choose as many interests as he wants (list generally varies between 15 to 25 interests, again depending on the firm). So we will have a matrix with values for all the resources for each user and the interests they opted for.

Table 1 Matrix of Users and their entered values for the resources and their interests

USER	R1	R2	R3	R4	R5	R6	R7	INTERESTS
U01								
U02								
U03								
U04								
U05								

Also will have another matrix, which will have *ratings* for all the Franchises that were recommended to the client **(on the scale of 5)**. So once we will recommend the franchises to our clients, they will be asked to rate them in rank them. Assuming that in our system, we will be recommending only 10 Franchises (F1, F2, ...F10) to our clients and that too in the order of priority.

Table 2 Matrix of Users and their ratings for the Recommended Franchise

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
U01										
U02										
U03										
U04										
U05										

Firstly, for any new user, we will use collaborative filtering (User-User Algorithm) to find out all the users who have similar values for the resources. So K-neighboring users with similar resource-values will be selected and then will use Pearson Correlation to find the closest User.

Secondly, Based on the level of similarity, a resource coefficient (r , $0 \leq r \leq 1$) will be calculated (with $r=0$ for zero similarity in the resource values and $r=1$ for 100% similarity in the resource values).

Thirdly, will find the weighted average of rating provided by each user to the recommended Franchises to them and the resource coefficient calculated in the previous step.

Example

Say we found out 3 users (U1, U2, U3) who have similar ratings as the new user X. based on the level of similarity we found out their resource coefficient (r_1, r_2, r_3). We found that $r_1 > r_2 > r_3$.

This signifies that U1 has the most similar values for the resources to X, and U3 has the least similar values. **This step helps us in knowing the importance or impact that each user can have on the new user.**

So now we will find out the weighted average for each of the Franchises that were recommended to the K-neighboring similar users.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
U01	3	5	2	1.4	3	0	0	0	1	4
U02	1.2	3	5	2	4	0	0	0	.5	1
U03	2	3	0	5	2	1	.5	0	0	0
X										

Like for F1, the value will be = $[(r_1 * 3) + (r_2 * 1.2) + (r_3 * 2)] / 3$

So the general formula will be,

$$= \Sigma \{(\text{Franchise Rating} * r) / \text{total number of users who rated that franchise}\}$$

Similarly we will calculate for other Franchises too and thus will get a single value for each of the Franchise. Now, on the basis of that we will get a priority list of all the Franchises that are relevant to the client (higher the value is, it is more relevant to the client)

Now we will denormalize those values and store them in the matrix as 'system rating'. It will be used later to evaluate our recommender system, by comparing system rating with the ratings provided by the user.

Working of Recommender System

Assumptions made

In this system we have assumed that initially when this recommender system will be incorporated, the consultancy firm need to

- Enter the required values of resources for each Franchise as specified by the corresponding Franchisor.

- Enter the sample rating for each of the Franchise based on statistical data received from market research, surveys, past experience and domain knowledge.

Step-by-Step scenario

Step 1: Each user provides value for the various resources and his interests

Step 2: All these values are normalized and stored in a matrix. Along with the normalized values, his preferences for various interests are also stored.

Step 3: Each user also rates Franchise that were recommended to them

Step 4: When new user enters, he provides his values and preferences.

Step 5: These values for the resources that new user has provided are normalized.

Step 6: By using Pearson Correlation, we find the neighbor users whose values are closest to this user's values

Step 7: Now we have identified the user closest to this user. Now we get the whole list of the Franchises that were recommended to similar users, and also rating for those Franchises.

Step 8: The next step would be to rank these whole set of Franchises.

Step 9: We rank this by finding the weighted average of user's rating of a Franchise and the resource coefficient (r) for that user.

Step 10: Hence by the end of this step we get one value for each Franchise. This value signifies the potential extent of match with the new user.

Step 11: As interests can't be normalized, so once the list of Franchise are given by the system, we should use a filter to check how many interests of this user matches with each of the Franchise in the list given by the system. The consultancy owner can set the threshold of interest match that minimum Z interests should match.

Step 12: So depending on the interest match we will get a final prioritized list of Franchises. **Its output is the list of final recommendations by the system**

Online/offline Computation:

Here we can see that Step 1, 2, and 3 will be executed offline i.e. system will be ready with the normalized values for the resources and ratings for the Franchise, for each user.

Advantages and Disadvantages

Advantages of Two-tier Algorithm:

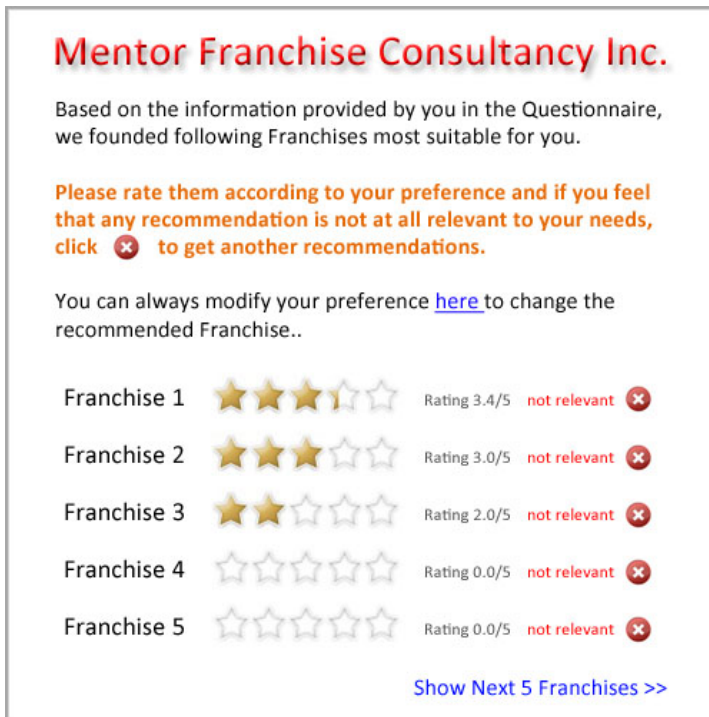
- In this algorithm we are considering both, the information provided by the user (about the resources) and his ratings for the Franchises.
- As most of it is executed offline (normalization of the matrices), so it's computed performance is high.
- If any user provides different rating then the system rating, the algorithm automatically takes that into consideration (through weighted ratings).

- For those **Franchises that haven't been recommended yet**, they will have values for the resources as set by the Franchisor and the default ratings will be entered from the market research and survey. So in that case, each non-recommended Franchise will act as a user and will be compared with the new user. So none of the Franchises will be left uncovered by the algorithm and all will have equal chances.

Disadvantages of Two-tier Algorithm:

- Since we are giving equal weight age to all the resources, so there might be a chance it might not be true. As in real world, the investment might be given more weigh than the time he can put into that business. So depending on the statistical data from the market research, we can provide weight to each of the resource.
- Any client is going to rate hardly 5 to 7 Franchises at a time, which is very sparse.

Interface



The interface that will be shown to the client will be as shown in Table 3.

Thus here when the client will select **not interested** for any of the Franchise, the next franchise from the list (not presently displayed in the UI) will be shown there.

Here the user can rate the Franchises on the scale of 5 by selecting stars.

Table 3 User Interface for the Recommender System

Evaluation

As we are recommending items which could become future businesses of our client and it involves high monetary risk, so we need very high precision. Our client's future is dependent on our recommendation. We are not interested in providing more options (recall) rather we are interested in providing less and accurate option (thus reducing our clients confusion).

So we will be using MSE (mean square error) to evaluate our system and thus punishing our system for bigger mistakes. We are more interested in the extreme of the errors and

thus not using MAE (Mean Absolute Error) as it focuses on the frequency of the errors. As the list generated by the system and the rating provided by the client won't be the final selection step of the Franchise. Consultant will be involved with the client to even further discuss the prospects of each of the recommended Franchises. So we just want to make sure that the list we get is very precise and accurate.

Privacy Issues

To protect our system from any kind of attack, we will use mobile-based confirmation system. As soon as client will fill the questionnaire and try to submit it, we will ask him to enter his mobile number. So a confirmation code will be sent to his phone. He needs to enter that to submit his questionnaire to us. This will prevent any manipulation from the client side, as user won't buy 1000's of mobile numbers to temper our result.

Again, once the recommendation will be generated, our consultant will give a call to the client and will personally inform him. If he finds that customer is authentic, then he will ask the client to go online and view the salient features of the recommended Franchise and rate them. Thus it's a double secured system that makes sure that the user is authentic and thus avoids any manipulation with the recommender system.

Conclusion

Having a recommender system for the businesses that involves future, money and trust of the client, is very crucial. If the Recommender system is not correct and our client is not satisfied with any of the recommendations made, then it's a big loss for the firm. So it is very important to make our recommendation very precise and accurate. The aim shouldn't be to provide a long list of options instead it should be to give them very limited but powerful options.

In Franchise Consultancy Business, Recommender System shouldn't be the only factor to decide that which Franchise will fit best for our client. Human knowledge and experience should be used to judge the recommended Franchise before finally opting that option.

In our recommender system as we are assigning resource coefficient (r) to the users based on their level of similarity, thus we are avoiding the situation where multiple fake profiles will be made to increase ratings of a specific Franchise. Again the weighted average of r and past Franchise rating takes into consideration any error made by the recommender system.

Reference

¹ www.thefreedictionary.com