RecIng – Analyzing the trends in Recipe and Ingredients used by Bohra Community

Syed Saqib Abbas Rizvi, Maaz Anzar Farooqui, Adil Saeed

Department of Computer Science and Information Technology

N.E.D University of Engineering and Technology. Karachi,Pakistan

[saqib4cs@gmail.com](mailto:saqib4cs@gmail.com), [maazanzar@gmail.com](mailto:maazanzar@gmail.com), engr\_ak[@outlook.com](mailto:adil.saeed@gmail.com)

**Abstract—** In this era where cost of everything is drastically increasing and everyone is demanding some changes in food industry so it is really important to know how ingredients or recipes are playing vital role in running the food industry, analysis of both two can give a good idea for an investor. In this regard we are going to do some analysis on Food recipes of a community that is named as ‘Bohra Community’.

***Index Terms*— recipe generation; ingredient cost; network analysis;**

# INTRODUCTION[[1]](#footnote-1)

E

very successful food business depends on having eye on your recipes and providing the best taste to your customers. Serving the good taste with reasonable price automatically expands the business. The barriers here to know about the both is not having the deep analysis of serving recipes and their cooking cost. If some knows the combination of recipes/ingredients those are in demand he could introduce more recipes and can server a unique taste. Network analysis of recipes/ingredients can give a clear picture of the trends those are favorite in customer. Alternatives of mostly used ingredients can be identified to make sure the 100% availability of the food. Cost cutting can be carried out after knowing the recipes and are not in demand.

# RELATED WORK

We explored some recipes generation app/stuff and found most the work is done around recommending the recipes based on the available ingredient, no one did analysis of ingredients, what relation they hold among different recipes? , are they active participant in recipes ?, what if they are not available at moment ?

# DATA

To have those analysis we were required to have some data. For the data we firstly visited the online data providers and got some open-source data, but that was not the desired data to perform analysis, some data sets are attached.

* <https://www.kaggle.com/shuyangli94/food-com-recipes-and-user-interactions>
* <https://www.kaggle.com/kaggle/recipe-ingredients-dataset>

After looking the above datasets we decided to go with any community based data set, community has own taste, has own culture, so knowing their culture in food would be fruitful and after that we Recipe data set of Bohra Community

# Analysis

We have Analyzed many networks related techniques to get most out network sciences. We have analyzed those following methods.

* **Closeness**

Nodes distance from a specific vertex to any other vertex in a graph is closeness.

The application of closeness is finding a place where maximum people can reach market.

Other example of this is to find the people (nodes) which can transfer information to entire community (graph) is a shortest time. This information can be anything like fake news or marketing or an election campaign

In our research closeness show the commonly used ingredient which is useful to maintain inventory so the restaurant maintains availability of meal

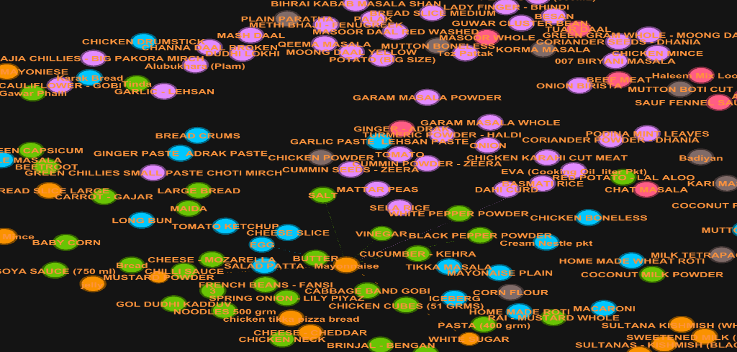


Figure 1 ingredients are commonly used in all recipes

* **HIT (Hub and Authority)**

It is a linked analysis algorithm proposed by **Jon Kleinberg.** Main purpose of this algorithm is to rate web pages. This algorithm uses the concept of Hub and Authorities.

* 1. **Hub:** hub is considered is the web page that contain large directories that is not authorize the information that they held. So, we can a hub is the source of pointer on other page.

We have used this concept to point out ingredients of meals. Each node show hub(ingredient) and size of node is representing the importance of ingredient.



Figure 2 Node show the ingredient and size of node represent its importance

**1.2 Authority:** The term authority is used to refer the web pages which are linked many different hubs.

In our research we use this method for identify the meals. so that we can analyze which ingredient are used for which meal. It helps us to identify those meal which have use max important ingredient. so, using authority, we have easily identified the cost of recipe which help us to make our business more profitable.

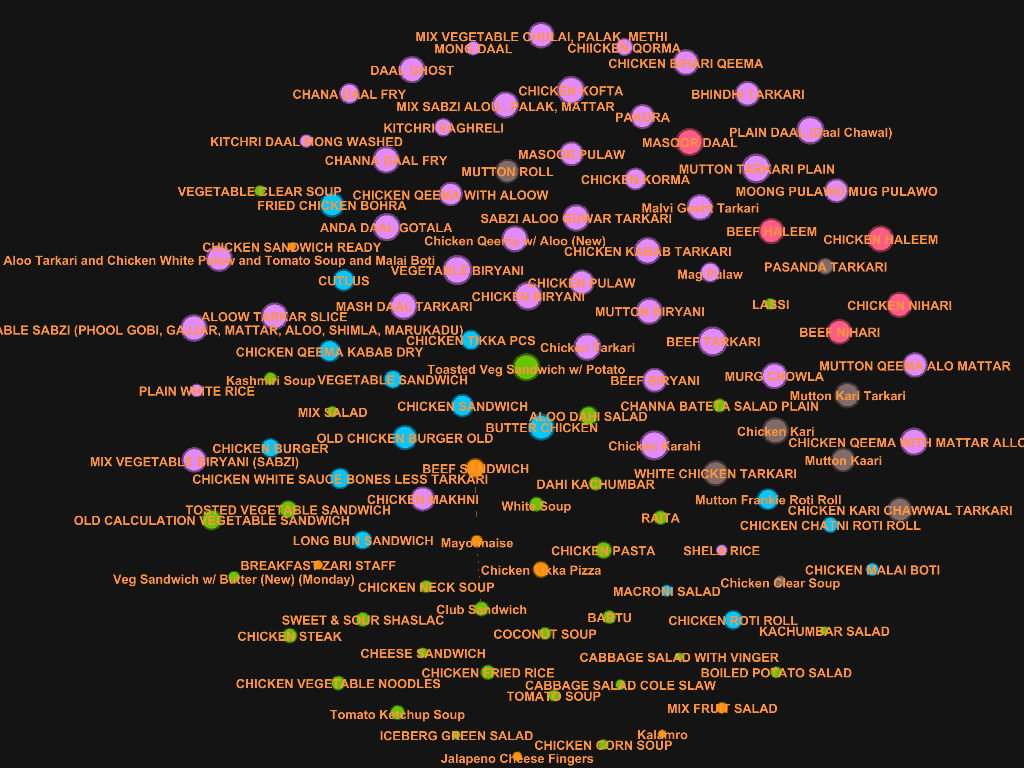
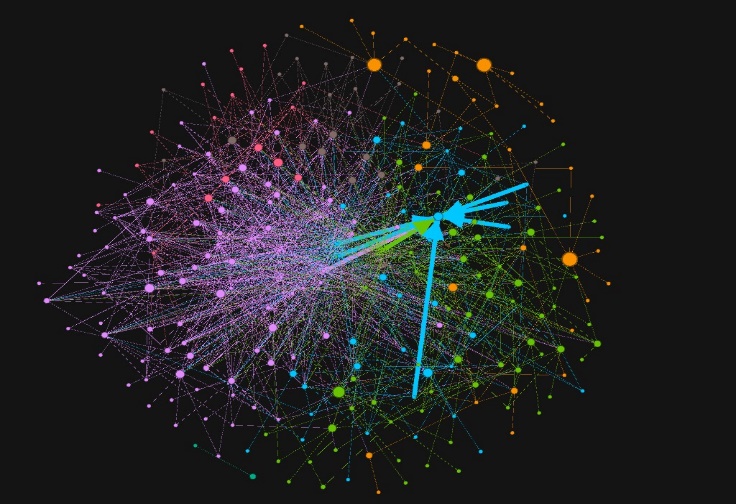


Figure 3: meals and size of node show the importance of meal

* **Page Rank**

PageRank Algorithm is one of the well know algorithm for numerical ranking of web page. Using this algorithm, we have identified relative importance of web page.

We have used Page Rank algorithm which is also identify the difference between ingredient and meals. Page Rank provide same and constant score of all ingredient since all ingredient have outdegree towards meals and assign different score to meals.



# Results

We find that Hub Represents the importance of nodes with respect to inventory.

* EVA (Cooking Oil liter Pkt)
* SALT
* CUMMIN SEEDS – ZEERA
* GARLIC PASTE LEHSAN PASTE
* GINGER PASTE ADRAK PASTE
* CHICKEN POWDER
* CUMMIN POWDER – ZEERA
* CORIANDER POWDER – DHANIA
* TURMERIC POWDER – HALDI
* TOMATO

# Conclusion

As shown in the above results we can see the important ingredients being used by Bohra Community. If someone is planning to start food business considering this community then he must have to consider the availability of important ingredients and cost.

Furthermore, in technical terms HITs algorithm is being used instead of Page Rank because Page Rank does not provide ranking of Ingredients while Hub score provides ranking of Hub based on Authorities.

References

W. Min, B. Bao, S. Mei, Y. Zhu, Y. Rui and S. Jiang, "You Are What You Eat: Exploring Rich Recipe Information for Cross-Region Food Analysis," in IEEE Transactions on Multimedia, vol. 20, no. 4, pp. 950-964, April 2018, doi: 10.1109/TMM.2017.2759499.

Nyati, U., Rawat, S., Gupta, D. *et al.* Characterize ingredient network for recipe suggestion. *Int. j. inf. tecnol.* (2019). <https://doi.org/10.1007/s41870-019-00277-y>

Hu Y, Zhang J, Bai X, Yu S, Yang Z (2016) Influence analysis of Github repositories. SpringerPlus 5(1):1268

Ahnert, S.E. Network analysis and data mining in food science: the emergence of computational gastronomy. *Flavour* **2,**4 (2013). <https://doi.org/10.1186/2044-7248-2-4>

Herranz, Luis & Min, Weiqing & Jiang, Shuqiang. (2018). Food recognition and recipe analysis: integrating visual content, context and external knowledge.

Ahn YY, Ahnert SE, Bagrow JP, Barabási AL (2011) Flavor network and the principles of food pairing. Sci Rep 1:196

Harvey M, Elsweiler D (2015) Automated recommendation of healthy, personalised meal plans. In: Proceedings of the 9th ACM conference on recommender systems. ACM, New York, pp 327–328

1. Special thanks to Dr. Qasim pasta for encouraging us to write this paper, he helped us throughout the whole process. We have learned alot about data collection to data analysis and then its visualization.

   Authors: Syed Saqib Abas Rizvi, Maaz Anzar, Adil Saeed  
   This is our collaborative effort. We have taken help from past research papers, course materials and network analysis techniques. [↑](#footnote-ref-1)