Friend Suggestion using Graph

Manishka Gautam, Himanshu Sharma, Mayank Chauhan, Himanshu Shukla, Nidhi Tawra

Abstract: ***The proposed design will present a new recommendation system for alumni association, which suggests friends to users based on their location and skills. The alumni can thus connect quickly and effectively with their peers. The project aims to connect the students, alumni and faculty through a single network where they can thus connect and collaborate for further ventures. The system provides the platform for students to gain momentum with respect to their career. The project aims towards giving a platform to the college authority which does not only serve the purpose of displaying college activities but is a full-fledged social network which includes suggestions based on graph database. It does not use any machine learning or data science algorithm but is purely algorithmic in approach***

Keywords: Graph Database, Neo4j, Suggestions, Web Application.

I. INTRODUCTION

The [10] existing alumni networking services are based on a casual web application which just enable the users to connect with their peers or just with the college at large. The following project will provide user-friendly environment with the recommendation system. This system will allow the user to find their batch-mates easily and thus connect with them as they would through a social network. The proposed design will implement on web-based application.

The results will show that the recommendations accurately return the friend suggestions.[1]Social networking services (SNS) provide a platform where people from different cultures interact and express their respective opinions similarly this system will provide the students, faculty and alumni to connect and have a common ground to use it as per their interests. This system can be deployed as a standalone app on Smartphone or as an add-on to existing social networks. The social networking apps and websites have great market in the industry as well. Apart from this the following project is economically viable as well.

Now since the current alumni systems do not include the recommendation system therefore this will be an edge over the rest of the systems. Since the system uses graph database, an important factor behind excessive usage of graph database is the increasing gains by leveraging [11] connected intelligence. Network management, relationships which include classifications and other related works have prominent usage of graphs and its algorithms.

Revised Manuscript Received on March 10, 2020.

Manishka Gautam*, Department of CSE, Meerut Institute of Engineering and Technology, Meerut, India.
E-mail: manishka.gautam.cs.2016@miet.ac.in

Himanshu Sharma, Department of CSE, Meerut Institute of Engineering and Technology, Meerut, India.
E-mail: himanshu.sharma.cs.2016@miet.ac.in

Mayank Chauhan, Department of CSE, Meerut Institute of Engineering and Technology, Meerut, India.
E-mail: mayank.a.chuahan.cs.2016@miet.ac.in

Himanshu Shukla, Department of CSE, Meerut Institute of Engineering and Technology, Meerut, India.
E-mail: himanshu.shukla.cs.2016@miet.ac.in

Nidhi Tawra, Department of CSE, Meerut Institute of Engineering and Technology, Meerut, India. E-mail: nidhi.tawra@miet.ac.in

II. LITERATURE REVIEW & RELATED WORK

With rapid increase in urbanization and industrialization there was need to improve innovativeness in technology as well. Earlier where people used to spend hours in travelling to the market getting stuck in traffic jams, nowadays it has become so easy due to commencement of online shopping. The E-Commerce market increased drastically with developments in their recommendation systems and suggestions. Facebook, had been showing friend suggestions since long time now but recommendation on online shopping sites has ease out our problems.

Where, once we used to ask our friends and family and take their suggestions, now the sites are providing us based on latest rating, user-experience, etc. Hence, technology has improved to great extent for good of common man. These recommendation systems can be based on collaborative filtering or content-based filtering. Content based filtering [2] compares dataset based on content which is read by users maximum. Collaborative filtering [3] is based entirely on user's behaviour. However in the following project the suggestions are based on graph algorithms Jaccard similarity. The reason for using graphs as the data structure was, that graphs provide results faster and are more efficient in comparison to other complex mechanisms. The suggestions and search shown in the project makes it more engaging since alumni portals are devoid of such features. Graph databases have manifold usages in everyday life as well. Human brain is one of the prominent example of the graph database where the nodes and edges are interconnected in the form of neurons and cells. Geography/cartography is also one of the field where graph plays vital role.

III. SUPPORTED TECHNOLOGIES

The project is a web based application which uses Node.js and React.js for entire development. To specify the technology used behind the project, Node.js is an open source [5], cross-platform runtime environment for developing server-side and networking applications. React.js is a JavaScript library which is used to build interactive user interfaces. The entire project is thus a JavaScript oriented product. In order to depict and provide outcomes related to the graph suggestions we have used Neo4j which is one of the efficient graph based data management system. Graph databases are an assistance in finding relationships amidst the provided data and thereby extracting their true value and one of the best and reliable service is Neo4j [8] which is implemented in Java. Due to Neo4j we were able to convert the semi-structured or unstructured data into an organised format (connected data).

IV. DESCRIPTION OF MAIN ALGORITHM

In order to create and provide graph based suggestions [8], the basic elements of any graph include the node and an edge. The node here represents the profile of the alumni or any individual who has registered himself/herself on the web-based application. The edge carries the weight of each path respectively. The weight is based on location and skills of the individuals involved. To explain it further we can assume x, y to be two registered users who are friends on the platform. Let "a" be the array which stores the skills of x namely: [html, css, js, python] and "b" be the array which stores the skills of y namely: [java, ruby, python]. Based on comparing the two arrays they have python as their common skill which eventually signifies that the weight can be 1 (here weight gets updated in proportion to number of skills in common). Similarly if location is found to be same the weight gets updated by 2.

The search algorithm used here is Breadth First Search with level 2 which searches subsequently the nodes and on finding those that have common grounds vividly depicts them as suggestion to the users. The pictorial representation of graph will clearly explain how the structure is formed for friend suggestions:

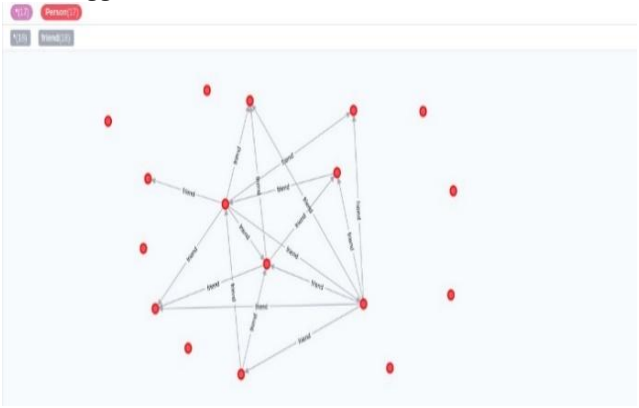


Fig 1: Graph Representation

V. DESCRIPTION OF MODULES

The following project being a web application is divided in modules which will be implemented synchronously. The initialization is conducted with student and alumni connect model which includes the registration module. This module will enable the novice to register on this web application. The admin module has further sub modules of sending invite to alumni, updating student details, search user details, view user details, maintain user details and post update information.

This is followed by the addition of event management and post modules. The post module will enable the admin, alumni and faculty members to post jobs and upcoming or latest college events. Event management module includes managing the meetups, events and alumni as well as college's extra-curricular activities related to alumni.

The chat modules will inculcate direct messaging amongst users with restrictions. This is the module which will give the user a touch of social media platform though it is not. The most important component of this modular structure forms the graph based suggestions which reside within suggestion module. This is the entire structure of the system comprising the above explained modules.

VI. ALGORITHM

The algorithm of the proposed project is explained below which includes variables User1 and User2 as the users of the web based platform that they are accessing after registering in it. Basically, in this algorithm we are using the Jaccard similarity to calculate the weight label for the relationship so that based on weight value we can provide the best possible path to reach the desired outcome. Graphs provide the best way to search and traverse since they are time efficient and space efficient as well. Profile1 and Profile2 are the profiles of the users User1 and User2 respectively who have acquired some skills. BFS up to level 2 is used in this algorithm to avoid the extra processing and wastage of resources. The highest weight calculated will be suggested as friends to others.

- User1 and User2 connect on the platform.
- Profile1 have some skills and Profile2 have some skills.
- Calculate weight label for relationship using Jaccard [4] Similarity:

$$V(A, B) = |A \cap B| / (|A \cap B| + |A-B| + |B-A|) \quad (1)$$

- Do traverse graph for User1 and find path (up to 2nd level i.e. friend of friend using BFS) between all users excluding current friends.
- Suggest top N friends with highest weight, where N is no of friends.

VII. RESULTS

The following table given below represents the dataset taken as a sample for the analysis and suggestions using graph database. The dataset is in the form of objects which represents the skills of users which are connecting and will thus show the path between them. The dataset includes three columns which are as follows.

- friends.name: This is the name of the user who will be connected to another user through an edge
- friends.weight: This is the weight of the edge calculated according to the algorithm
- skills: These are the skills of respective users.

Table I: Skill Set

"users"
{"skills":["scala", "java", "python", "html"], "name": "User2"}
{"skills":["java", "c"], "name": "User3"}
{"skills":["c++", "c", "python", "java"], "name": "User4"}
{"skills":["java", "python", "c", "html"], "name": "User1"}
{"skills":["c++", "c", "python", "java", "scala", "html"], "name": "User5"}

The following table includes the relationship between the nodes which is well depicted in the graph being formed as a result. The three columns have the values of weights carried by the edge and the other two columns represent the users which are being connected using the edges.

Table II: Relation between Users

"friend.name"	"friend.weight"	"friend.name"
"User2"	3.3	"User5"
"User3"	2.5	"User4"
"User3"	1.6	"User5"
"User1"	3	"User4"
"User1"	3	"User2"

The table shown below is the end result and implementation of algorithm:

Table III: Final Outcome

users	before	after
User1	User3	User5
User2	User3	User4
User3	User2	User1
User4	User2	User2
User5	User4	User1

The below shown image is the pictorial representation of the graph formed after applying the algorithm:

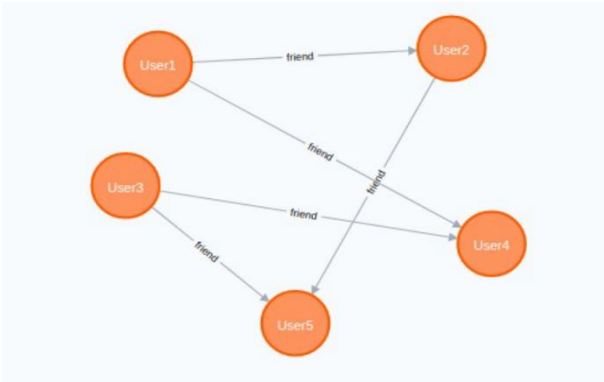


Fig 2: Final Result in Graph Form

VIII. CONCLUSION

In this approach we presented the design and implementation of an alumni network which provides an environment of social network while protecting its age-old usage of connecting the college authorities with their alumni. The graph based suggestion has been proved to be more efficient since it takes less time for traversal and searching. Apart from it SQL database or traditional database has defined columns and rows where data is stored in the form of tables which makes data retrieval quite hectic and less efficient. Therefore, a graph database might be described as more "relational" than a relational database. Graph databases shine when the goal is to capture complex relationships in vast webs of information [6].With the help of graph database we concluded that if we desire real-time, efficient insight and action then graph database is the best choice for such projects. Above all graph database transforms the widespread data which is in relational or rather traditional format, is converted to meaningful relationships which are far easy to understand and draw outcome from. Neo4j provides high availability with transactional guarantees also. Thus the project had enables us to draw and derive meaningful insights and therefore we could pen it down in the form of research paper.

REFERENCES

1. SNS,https://www.tandfonline.com/doi/abs/10.1080/10447318.2012.756333
2. Content based filtering, https://towardsdatascience.com/brief-on-recommender-systems-b86a1068a4dd
3. Collaborative filtering, https://towardsdatascience.com/brief-on-recommender-systems-b86a1068a4dd
4. Jaccard similarity, https://www.statisticshowto.datasciencecentral.com/jaccard-index/
5. "Node.js https://www.tutorialspoint.com/nodejs/nodejs_introduction.htm
6. "relational" https://www.infoworld.com/article/3263764/what-is-a-graph-database-a-better-way-to-store-connected-data.html
7. Breadth First Search, https://www.geeksforgeeks.org/breadth-first-search-or-bfs-for-a-graph/
8. Neo4j,https://www.qhmit.com/neo4j/tutorial/ https://www.researchgate.net/publication/317349295_Calculating_the_User-item_Similarity_using_Pearson's_and_Cosine_Correlation_Senthilkumar_M
9. "alumni", https://www.redbrickresearch.com/2015/10/29/the-importance-of-the-alumni-network/
10. "leveraging", https://neo4j.com/news/the-importance-of-graphconnect-first-graph-database-conference-ever/

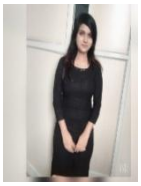
AUTHORS PROFILE

Manishka Gautam, is pursuing B.Tech. in Computer Science and Engineering from Meerut Institute of Engineering and Technology, Meerut and is graduating in the year 2020.

Himanshu Sharma, is pursuing B.Tech. in Computer Science and Engineering from Meerut Institute of Engineering and Technology, Meerut and is graduating in the year 2020

Mayank Chauhan, is pursuing B.Tech. in Computer Science and Engineering from Meerut Institute of Engineering and Technology, Meerut and is graduating in the year 2020

Himanshu Shukla, is pursuing B.Tech. in Computer Science and Engineering from Meerut Institute of Engineering and Technology, Meerut and is graduating in the year 2020



Nidhi Tawra, pursued M.Tech in Computer Science and Engineering from JAYPEE Institute of Information and Technology, Noida in 2017 and B.Tech in Computer Science and Engineering from Krishna Engineering College, Ghaziabad in 2014. Currently she is working as an assistant professor in Meerut Institute of Engineering and Technology and is also the mentor/guide of this project