Spring 2016

Hybrid Algorithm for Matching Profiles and Social Networks

Peter Elliot Garcia
Bard College

Hybrid Algorithm

for Matching Profiles and

Social Networks

A Senior Project Submitted to
The Division of Science, Mathematics, and Computing
of Bard College

by
Peter Elliot Garcia

Annandale-on-Hudson, New York

May 2016

Abstract

The online dating experience has revolutionized the way how many of us form relationships with one another. It turns out that many people struggle with finding their perfect match because their self interest tends to get in the way. They may find someone of their liking but someone else might spark their interest so this project aims to eliminate this uncertainty by creating a hybrid algorithm to match people within a dating pool. The algorithm acts as a self-enforcer to create the best match possible under the guise of a social network application. The application will act as a probe to gather the information from the users. The results will help us examine the best ways to match people up in the online dating world.

Dedication and Acknowledgments

This project is dedicated to my family for their continuing support and love. I owe a special thanks to my mother who sacrificed so much for me to be in the position I am in today. Without her love and encouragement I'd probably be lost in the streets of the Bronx. I'd also like to thank my sister Michelle who helped raise me alongside my mother. Without you by my side, I'd probably be a douchebag. I'd like to thank both of my brothers, Steven and Wesley, for their love and support. Shout outs to my beautiful baby nephew, Liam, for bringing pure joy and love in to my life. I want to thank my advisor, Robert McGrail, for his support throughout my senior project journey. Without your help this project would not be possible. Mary Ann Krisa for kidnapping me during L&T. I would have transferred out from Bard if it wasn't for you. Shout outs to all my friends, Troy, Noah, Sadiki, and Isaiah, who have been with me through the hardest of days. You're all kings but also my children. And last but not least, to all the black men and women who lost their lives due to the racist, corrupt police violence in this country. Your names and lives will never be forgotten.

Table of Contents

1
Introduction

1.1 *Food and Socialization*

When people gather together across various cultures to engage in one activity, it creates a commonality between those cultures. A culture is defined as a way of life of a group of people--the behaviors, beliefs, values, and symbols that they accept, generally without thinking about them, and that are passed along by communication and imitation from one generation to the next. We can consider mealtimes as a cultural site that everyone shares regardless of location. These cultural sites are used as a device to create competent and appropriate members of society. The socialization that occurs may differ in the specificities between different cultures but in broader terms, they are all quite similar. Cultural sites, as defined by Elinor Ochs and Merav Shohet, are conceptualized and historically durable yet transformable, socially organized and organizing, and tempo-spatially situated arenas, which are laden with symbolic

meanings and mediated by material artifacts[1]. These sites are created as such because of the constant participation that occurs from the participants. People are expected to behave in a certain manner when they come together at these cultural sites because of the social implications that these sites have. They may not be aware of this particular behavior, but the settings in which they choose to engage in dictates the behavior amongst the participants involved.

With this notion of cultural sites, we can assume that mealtimes can be regarded as locations for social understandings of the world. Mealtimes are used to drive the culture forward through social capacities. It is because of this that mealtimes constitute universal occasions for members not only to engage in the activities of feeding and eating but also to forge relationships or build stronger ones. The relationships that are created within these locations are specific to mealtimes. It is generally expected for people to behave in a manner that is associated with mealtimes. This manner is developed from a long history of repetition associated with mealtimes. Since childhood, children are socialized into the importance of food sharing in the family unit through accompanying others in continuous cross-household visits that involve preparing, offering, and consuming food.[2] Food is viewed as symbolic, with its members who imbue particular kinds of qualities of food with sentimental, moral, religious, and health-related meanings. Since this is established in the home, it permeates to how people interact with one another where mealtimes are not happening in the home but rather in restaurants.

---

[1] Elinor Ochs, Merav Shohet, *The Cultural Structuring of Mealtime Socialization* (California: UCLA, 2006), 35-37.

[2] Ochs, *The Cultural Structuring of Mealtime Socialization, 37.*

The socialization between people has evolved overtime. In the past, people were more social on an interpersonal basis; however, with the advancements made in technology and with the quick rise of the internet, the ways in which we communicate with one another has altered throughout the decades. With the increase of social networks and online media access, there have been many changes in how people interact with one another. Since the invention of the mobile phone, people have been able to stay in contact with instant access. Mobile phones have also created an incentive for people to stay in contact with each other through their applications. While the smart phones have provided great use of a portable phone, the applications that are available has produced new industry for people. This industry includes various platforms of social networks such as Twitter, Tinder, Instagram, etc. These platforms provided people to have access to one another, even when they are not in the same location. With all these advancements made in technology, people are also able to remain in contact with people they've known while also having the opportunity to meet others.

Social networks are special because they reveal a different type of human behavior that has never existed before. The ways in which people have shared information through social networks has altered the way we take in the details. It is because of this interaction that data collectors are able to gather information about people based on their communication online. Information that has been collected has revealed to us what online communication can tell us in terms of what is trending, how users feel about their social network, and the cultural attitude towards current events.

Although there are differences amongst every social media site, all social networks have three things in common: they allow people to communicate with one

another regardless of distance, they are all visually impressive, and they grant a user the ability to see who is within their network. Since social networking is all about being social, every platform grants their users the ability to talk to each other. It comes in many forms such as commenting on photos and/or videos on Facebook and Instagram, the ability to mention one another on Twitter, access to instant message each other, and more. By doing this, you are building connections with people you do not know and also creating stronger connections with the ones you already know. A social network must contain a form of communication, otherwise it would not be considered a social network.

A good social network needs to be aesthetically pleasing to the eye. It also needs to be simple enough for the user to use. This is crucial because if the platform is difficult to navigate or look at, this runs the risk of losing users. The ability to see those who are within your friend group is another critical component. Pictures, videos, and memes are all ways to keep people visually interested in a social network.

A social network is nothing without the network you create for yourself. Seeing who is within your network is an important feature for all users. You are able to see what you all have in common and this gives each user the key to learning about what is trending. The need to be technologically connected with one another has grown significantly with the release of smartphones. Smartphones, such as the iPhone or Android phones, have given their customers access to these social platforms. People have connected through photos, videos, music, and a variety of other things. It is through those networks that data collection has revealed what human interaction online can tell us.

1.2 *Mining Twitter*

Twitter is a great source to examine as it reveals much about people and its structure. Twitter is a rich source of social data that is great for social mining because of its openness to the public, broad appeal to any user from any walk of life, and clean well documented application program interface (API). Twitter is also interesting because tweets are occurring at the speed of thought and are available for the public eye at near real time, representing the broadest cross-section of society at an international level, and are so inherently multifaceted.[3] Tweets and Twitter are connecting people in a variety of ways, from the short dialogues between people and other things that their users actually care about.

Twitter has made it so that every user is heard, their curiosity can be satisfied, it is easy to use, and it is available anywhere and at any time. We are able to analyze trending topics that are occurring and tweets with frequency analysis. With all these tools that Twitter grants to us, we can observe what is being talked about the most and at what frequency. Twitter displays the true points of humanity. We have a deeply rooted need to share our ideas and experiences, which gives us the ability to connect with the people we know and don't know so that we can be heard, and feel a sense of worth and importance. We are curious about the world around us and how to organize and manipulate it, and we use communication to share our observations, ask questions, and engage with other people in meaningful dialogues about our quandaries.

---

[3] Matthew A. Russell, *Mining the Social Web.* (California: O'Reilly Press, Inc, 2013) Chapter 1

One way to describe Twitter is as a microblogging service that allows people to communicate with short, 140-character messages that roughly correspond to thoughts or ideas. In that regard, you could think of Twitter as being similar to a free, high-speed, global text-messaging service. In other words, it's a glorified piece of valuable infrastructure that enables rapid and easy communication. It doesn't adequately address our inherent curiosity and the value proposition that emerges when you have over 500 million curious people registered, with over 100 million of them actively engaging their curiosity on a regular monthly basis.[4]

The extra ingredient that makes it sufficient is that Twitter's asymmetric following model satisfies our curiosity. It is the asymmetric following model that casts Twitter as more of an interest graph than a social network, and the APIs that provide just enough of a framework for structure and self-organizing behavior to emerge from the chaos.[5]

Social websites such as Facebook and LinkedIn require the users to accept a friend request, whereas Twitter's relationship model allows you to keep up with the latest happenings of any other user, even though that other user may not choose to follow you back or even know that you exist. Twitter's following model is simple but exploits a fundamental aspect of what makes us human: our curiosity. Whether it be an infatuation with celebrity gossip, a keen interest in a particular political topic, or a desire to connect with someone new, Twitter provides you with boundless opportunities to satisfy your curiosity.[6]

---

[4] Russell, *Mining the Social Web*, Chapter 1

[5] Russell, *Mining the Social Web*, Chapter 1

[6] Russell, *Mining the Social Web*, Chapter 1

1.3 *Mining Emails*

Mail archives are arguably the ultimate kind of social web data and the basis of the earliest online social networks. Mail data is universal, and each message is inherently social, which involves conversations or dialogues with a minimum of two users. Furthermore, each message consists of human language data that's inherently expressive, and is laced with structured metadata fields that anchor the human language data in particular timespans and unambiguous identities.[7] We can introduce some fundamental tools and techniques for exploring mailboxes to answer questions such as:

- Who sends mail to whom (and how much/often)?
- Is there a particular time of the day (or day of the week) when the most mail chatter happens?
- Which people send the most messages to one another?
- What are the subjects of the liveliest discussion threads?

Although social media sites are racking up near-real-time social data, there is still the significant drawback that social networking data is centrally managed by a service provider that gets to create the rules about exactly how you can access it and what you can and can't do with it.[8] Mail archives, however, are decentralized and scattered across the Web in the form of rich mailing list discussions about a litany of topics, as

---

[7] Russell, *Mining the Social Web*, Chapter 6

[8] Russell, *Mining the Social Web*, Chapter 8

well as the many thousands of messages that people have tucked away in their own accounts.[9] Mining mail archives is one of the most essential capabilities to have.

1.4 Mining the Semantic Web

When we think of the semantic web, we must think of a version of the web that we already know, except that it has evolved to the point that machines are routinely extracting vast amounts of information that they are able to reason with and make automated decisions based on that information. Microformats are a key component to the semantic web. They are a relatively simple technique for embedding unambiguous structured data in web pages that machines can trivially parse out and use for various kinds of automated reasoning.[10] The implications of microformats for mining the web are vast.

Microformats are simply conventions for unambiguously embedding metadata into web-pages in an entirely value-added way.[11] Microformats are formatted to parse recipes online, determine a location of different users online, identifying people and companies, embedding résumés and much more. Although it is difficult to link microformats as social data, it is still interesting and plays an increased role in social mashups.

About.com is one of the more prevalent sites that uses microformats such as hRecipe. hRecipe is used for identifying recipes online and is widely used by niche food

---

[9] Russell, *Mining the Social Web*, Chapter 6

[10] Russell, *Mining the Social Web*, Chapter 8

[11] Russell, *Mining the Social Web*, Chapter 8

sites such as subdomains on about.com, for example, thaifood.about.com. We can use this to improve online dating. We can consider an online dating service that attempts to pair people together for dinner dates. One could reasonably expect that having enough access to ones geolocation and cuisine information linked to specific people would make a profound difference in the success rate of first dates. People can be paired together based on two criteria: how close they live to each other and what kinds of food they eat. For example, a dinner date between two people who are organic vegan eaters would go better than a date between a barbecue lover and a vegetarian. Cuisine preferences and whether specific types of allergens could affect a date.

1.5 *Application*

For the reasons presented above, building an application where people are matched together for their similar tastes in food would create a new online dating service. Like all dating sites that currently exist, an application will be able to give people an opportunity to connect with each other based on their location and food preferences. The mission of this app will be to bring singles together for a date they will find most enjoyable and ones in which they are looking for. Profiles will include one photo album that users can scroll through, as well as choosing your preferences in regards to the partner they are searching for. Members will be able to easily see who is around them and scroll through their network at ease. Since locations are needed to view who is in the nearby location, anonymity will not be taken into account for the sake of displaying who the users are. All of these aspects brought together creates a new style of online dating.

The application starts off by asking for the user to agree to share their location because, similarly to existing applications, a users location will determine who is in the surrounding area. Since people have a variety in tastes for food, the app will need to get a sense of what your personal preferences are. After a user agrees to share their location, each person will be asked to complete a questionnaire. The questionnaire contains questions that are not only essential to creating a network for a user, but to establish themselves for other users networks. There are a series of questions to answer for every user and based off of the answers given, there will be other users that are brought to your network if there are similarities detected. The questions that are posed to every user gets a range of their preferences. For example, questions will ask for a range of numbers, specifically from 0 to 5 for some questions and 0-10 for others. A user will be asked from a scale of 0-5, how much do they like meat, with 5 indicating that they very much prefer any kinds of meat and 0 implying that they are either a vegan or vegetarian. There are questions asking a user how they would scale their liking for coffee, tea, dairy, gluten, spice ratio, and alcohol. The reason for the questions simplicity is because we are trying to get a brief idea of your preferences so that users get a synopsis of you to spark a curiosity that can lead into having interest, with the possibility of a date coming about for the two users.

Based on these commonalities, users are able to interact with each other in a network full of people who have those similar interests. The homepage allows the user to be fully immersed within their network. They are immediately exposed to every person within their network. The design of the homepage enables the user to view as many people within their general network. For every user that a person would not be

able to view, the user can simply scroll to the left or right or whichever direction they would like to go in. This gives the user a more interactive experience rather than other social media platforms that only allow you to view people on an individual basis. Figure 1.1 demonstrates a mock layout of how the interface would be displayed for the users.

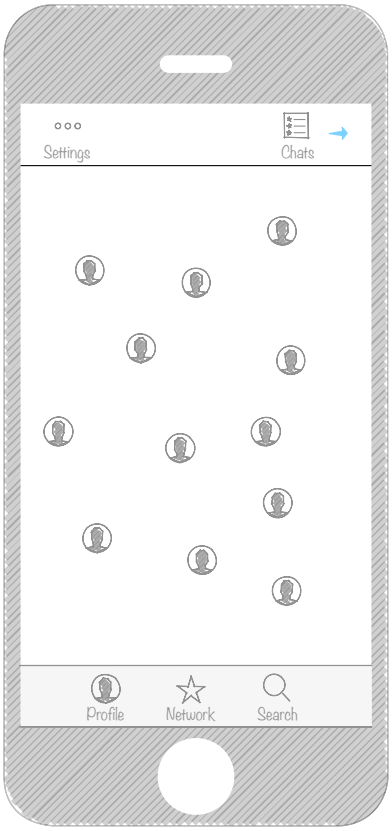


Figure 1.1 A mock screen display of the homepage.

If other users are signing up for this platform and after they complete the questionnaire, if their answers matches up with an existing users answers, the newly signed up users will be added to an existing users network. This grants anyone the capacity to expand their networks as much as possible, giving each individual the capability to meet more people. A user is also given access to change their answers from the questionnaire. Members are allowed to resubmit their answers from the questionnaire which permits everyone to create a new network for themselves. Once this is done, a new network is created for the user in order to find a new match.

Once a user is given the opportunity to see who is available within their network, they have the ability to select each person individually and view a blurb about them. This gives every user a chance to learn about one another but it limits the information that is displayed. This gives the user a chance to explore more about the person they are interested in. The blurb contains the similar food interests along with a brief statement about the person. A user is able to scroll through an album of photos that was uploaded to their profile. This synopsis of each users keeps the curiosity alive between each individual. Figure 1.2 is an example of the brief profile view.

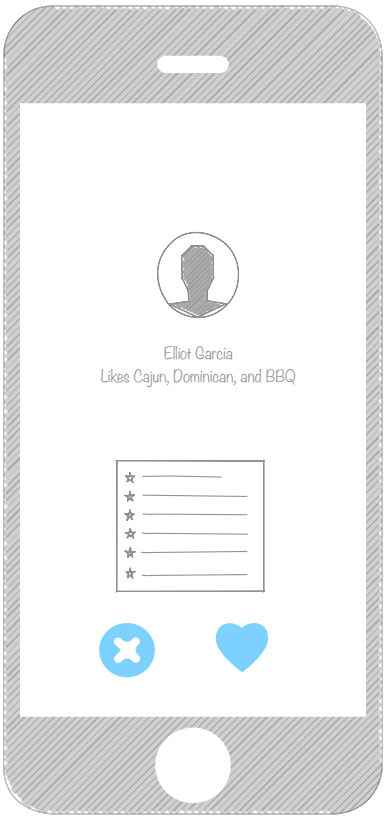


Figure 1.2 A mock screen display of the brief profile view.

Each user is granted a “like” and “dislike” selection for everyone within their network. When a person finds someone that they are willing to get to know and selects the “like” option, that user will then be highlighted within the person they are interested in network so that both of the users are made aware of the interest. When a user hits the “dislike”

option, the person within their network is now discarded and will no longer be available to view on the homepage. If person A has shown interest in person B, person A's profile would light up to indicated to person B that they have revealed an interest in them. An example of this can be seen in Figure 1.3.

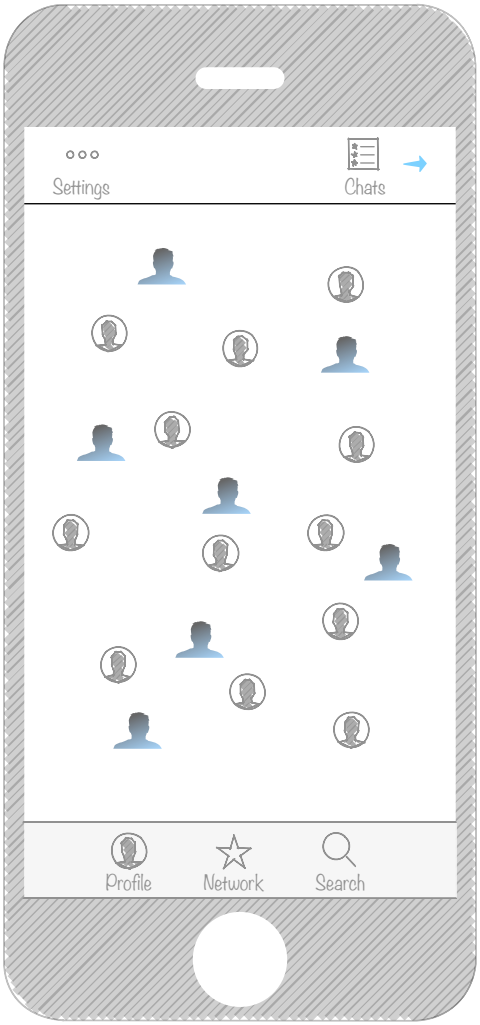


Figure 1.3 A mock screen display of the interested highlighted users in a network

After both users take interest into each other and agree to further their communication, a chatroom will be available for the conversation to continue. If the chat is going in a direction that leaves both user wanting more out of it, there will be lightning bolt button to create a date between the two users. View figure 1.4 to view a mock screen shot of this.

Figure 1.4 Mock screen display of the chatroom between two users.

On the backend database, the answers from the questionnaire combined with a set of restaurants will produce a cuisine venue for the users that are special to each matching. The restaurants provided are only accessible in the location that both users were found in. If it is unable to locate any restaurants in the area, it will present an alternative choice.

Social networks are a prime market to see what the consumer is asking for. This can be done in various ways for this project. Based off of how many people join the

application, we are be able to determine many things that will be gathered into the database. First, we are able to see which locations are more densely populated when using this app when every user agrees to share their location. Since this application will not be limited to the United States, we will be able to see which locations around the world are on this social network.

After we are able to see which locations are more popular, we can then determine which restaurants are being accepted to go to more than others. Are barbecue restaurants more popular than seafood restaurants? If so, which location is this occurring in? Which city has a greater success rate for matchings? How likely are matches occurring in Los Angeles compared to Chicago? Are dates more successful in New York City than in Boston? Giving people the incentive of dating one another, the data that is collected can reveal a lot about the success rate of dates when matched together with food across the globe. The use of the semantic web will provides a common framework that allows data to be shared and used across any boundaries. The information gathered can display what is trending in the food world.

There are also alternatives to sharing information other than the semantic web. Clustering memes in social media is also a common way for information to be passed along from person to person. It is through meme clustering that gives researchers the chance to study human social behavior while also challenging the capability to study the massive data streams. The challenge at hand is to see whether the various ways of activity are spontaneous forms of communication or engineered forms of misinformation campaigns.[12] Because of this, a formal definition of a meme is needed, which is the

---

[12] Emilio Ferrara, et al., *Clustering Memes in Social Media, http://arxiv.org/abs/1310.2665* (October 10, 2013).

process of how information is passed from person to person through social networks.[13] Once memes are detected, there are supervised learning methods that can be applied to classify the types of communication and information that is occurring in these social networks. This ties back to earlier in the chapter where we discussed mining Twitter and emails. The information available to us can reveal a lot about people.

Memes are not the only way of detecting information and classifying them. Tags and keywords are other forms to analyze the most popular types of communication that are arising in social media. Trending topics on Twitter and Facebook are perfect examples of this. Each of these social networks explicitly alert their users of the common topics of conversation within their networks. These memes, however, are not specific words that are highlighted the way tags and keywords are. Memes are a humorous images, videos, pieces of text, etc. that is copied (often with slight variations) and spread rapidly across the internet from the users. Memes, for example, have often been used to to criticize politics, the media, etc. but this is be done in a satirical way which is a way for memes to spread information at a faster speed.

The relevance of meme detection is that it reveals what information is being spread and seeing how people interact with one another given the circumstances that they are in. For example, under the specialized context of Facebook, anything and everything can be shared, talked about, etc. but if the circumstances were altered so that the users were under the guise of a controlled theme, would the information that is being talked about change? Would there be a reason to detect the topics of conversation if everything is indirectly geared to focus on a specific topic?

---

[13] Ferrara, *Clustering Memes in Social Media, http://arxiv.org/abs/1310.2665 (October 10, 2013).*

The purpose of this project is to explore the best possible way to match people together through food within a given dating pool. We will be using a hybrid algorithm, that we will discuss in chapter 3, that can eliminate the self-interest of a user within the same dating pool. The algorithm should eliminate the self-interest of a user and be able to match two users based on their answers from the questions, thus being a self-enforcing algorithm. The questions presented above are ways we can further explore how social networks, whether through mobile applications or a website, and seeing what can be concluded from the data that has been gathered. If we view social networks as objectively as possible, then it can shed light on a recent phenomenon and how humans interact on a platform that has been provided to us.

2
Existing Competition

2.1 User Interface of Applications

The application that is designed for this project has many similarities to existing applications but there are also many differences between each of them. The basis for the application at hand has incorporated many ideas from apps such as Tinder and match.com. But with these similarities, there are also differences between each platform that distinguishes itself for the users.

Navigation is essential in building apps because if the design of the application is too sophisticated or possibly too simple for the user, then the possibility of members leaving the social network may increase. Apple specifies the three important qualities for designing apps for the iOS platform: deference, clarity, and depth. The user interface

(UI) helps people to understand and interact with the content, but never competes with it when it comes to deference. Clarity deals with everything incorporated in the application, meaning, the text is legible at every size, icons are precise and lucid, and many other aspects that are factors for the clarity for the UI. The visual layers and realistic motion impart a vitality and heightens people delight and understanding for the depth of the application.

There are more features to consider when designing an application for iOS. First, a designer must look past the UI to the apps functionality and affirm its relevance. Secondly, designers must use the themes of iOS to inform the design of the user interface and the user experience. Finally, a designer must be sure to design the UI to adapt to various devices and modes so that users can enjoy an application in as many contexts as possible.

2.2 *Tinder*

Tinder, an app designed as a match maker based off of your location and physical appearance, has changed the way we interact with one another. The application will hone in on your location and will allow you to view who else has access to the application. It grants a user the ability to decide the mileage they seek to explore, whether it be five miles or twenty five miles away, to get a broader idea of who is around them. The application also only lets you login with an existing Facebook account. When a user signs in with their Facebook account, pictures and information are taken from their account and added to their Tinder profile. The information includes pictures and a brief synopsis about a person. If an Instagram account is linked to the Facebook

account, pictures from the Instagram account are now available to be viewed on the Tinder profile, giving each user more details about the person they are viewing. Tinder also displays common connections betweens each of the users. It compares your Facebook friends with those of your matches to display any common connections. It displays the distance in mileage between each user.

Upon logging in, the application asks if the user will allow their location to be used. If someone decides against this, then the purpose of this app will no longer be of use to that user. Each user is also allowed to receive notifications from the application. The notifications are used to alert each user when a match has been made or when a message has been received.

A user can anonymously “Like” or “Pass” on people Tinder suggests. A user will have to swipe right on a photo to like them or swipe left to pass. If a user has liked someone and they happen to like you back, then this is considered a match that has been created. Once this match occurs, you are able to chat with the person from within the app. The users are able to get to know each other more through this. There is also a feature where you can “Super-Like” a person who appears on the stack of cards. What this means is that when a user super-likes someone, they will appear on the other users stack of cards, highlighted in blue to alert them that they are very interested in them. The other user still has the option to deny them regardless of the super-like. Another interesting feature of Tinder’s is the retry option. What this does is give every user the option to rethink who they swiped right or left on. If there is someone you did not find attractive and accidentally swiped right to like them, you have the ability to bring that option back and change the decision you made and vice versa for anyone you may

have disliked. You can also change the age range of who you're looking for with the youngest age being 18 and the oldest being 55 and older. A sample page of tinder can be viewed in figure 2.1.

Figure 2.1 A sample display of a tinder page.

The catch to this, however, is that this feature is not accessible to everyone. The only way to access this feature is to pay for it, which is called Tinder Plus. Tinder plus includes features such as the rewind option, more super-likes to give out, the ability to not have any advertisements shown, and the option to change your location so that you can create matches around the world. Tinder plus is available for $9.99 for one month, $5.83 for six months, and $.58 for yearly subscriptions.

There are immediate similarities between both applications. Firstly, they're both central on the users location. If a user is denying access to their location, then the socialization between the individual and the network they've created is now unaccessible. The importance of this feature is crucial for this application. Without people to engage with then it loses the social weight to it. What it also shares with Tinder is the ability to view the distance to between both users.

With these similarities in mind, there are also differences between both applications. While location is crucial to meeting new people for your network, that is not the only source for determining who you can or can't encounter. The means of viewing the people within your network is also drastically different than how Tinder allows you to view their users. With a stack of cards style of display, you are only able to view Tinder users on an individual basis. The application for this project, however, you are able to view your network in an atlas style display. This gives each user the opportunity to not only see a larger number of people who are in their network, but they are able to navigate it in a way that is more user-friendly by scrolling around the page.

2.3 *Match.com*

There are other social networks that are more designated to creating relationships rather than providing potential "flings" such as Tinder does. Match.com is a network, amongst many other social networks, that is designed to assist single people on their quest for relationships and love. Founded in 1995, the purpose of this new style of dating, now known as online dating, created a chance for people to meet their special

one. This isn't a restriction to just the United States as match.com has also provided users from twenty four countries and territories to find love.

The method used to produce these relationships is rather simple: they give single people an opportunity to present who they are through a serious of free written response questions and multiple choice questions, while also giving a user the opportunity to upload photos of themselves with the limit being twenty six pictures. In addition to the photos that have been added, the users are asked about their selected preferences as to who they are searching for. Once this process is complete, users are able to scroll through various profiles, view photos, and read about potential matches in the area. Tinder allows you to view many user profiles with anonymity but once a match has been made, that anonymity no longer exists. You are able to view who are conversing with as stated beforehand. Match.com, however, firmly believes in the privacy and integrity of their users. The communication in place for everyone happens through an anonymous network. It is up to the user to give the approval of any information they wish to reveal. Up until that point, their name and contact information is kept private until they decide to reveal it to their match. A sample match.com page can be seen in figure 2.2.

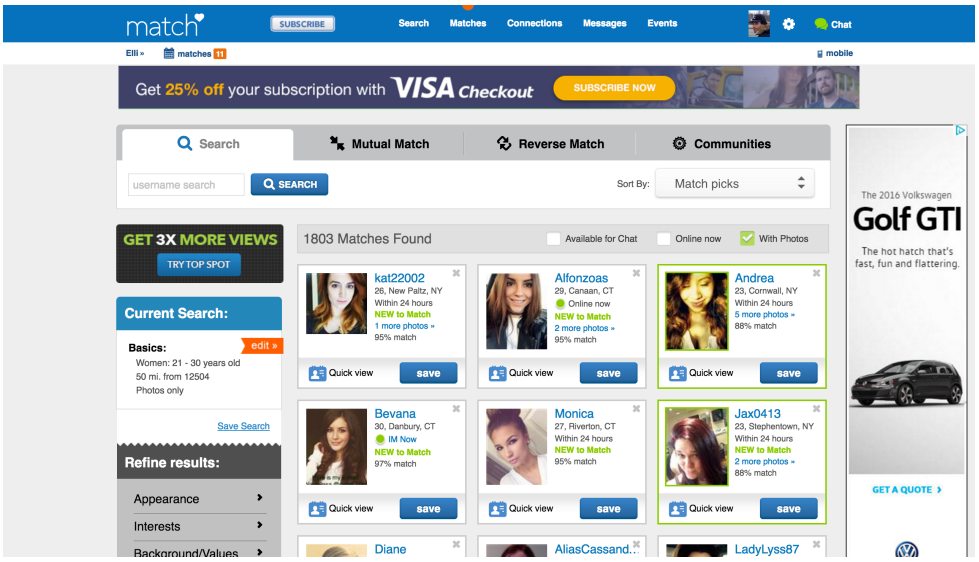


Figure 2.2 A sample homepage for match.com

The survey that is provided for the users to specify their preferences is really an in depth analysis to find the perfect match for a user. Questions are as such: “What’s her hair color?”, “What’s her eye color?”, “What’s her body type?” and much more. The specificities associated with the perfect match are quite interesting as it tries to do an in depth analysis of the singles interests. The questionnaire extends to analyze the user to once again get another in depth idea of who you are as a person. There is a stark contrast between the two analyses of your ideal person that you are searching for and the questions posed for yourself. The generic questions include “What is your height?” and “What is your hair color?” and so on. But as you make your way through, the questions take a shift to “What is your salary range?” and “How often do you exercise?”. There are sections for written responses from the user. The paragraph can be used for the user to either take the time to describe themselves or write about their potential match that they are searching for. Match.com also gives every user the opportunity by offering to share your profile to other online dating sites. By accepting this offer, your chances to meet someone will double because a users profile and information will be shared on the sister sites of match.com such as chemistry.com, thus increasing the chances to find your match. This is known as ‘Member Spotlight’. It also allows your profile to be highlighted. Once your profile is highlighted, the allotted time granted for every user is two weeks. Once a profile has been highlighted, a user cannot be removed until the time has expired. To be considered for a profile highlight, a user must have an attractive profile photo of them self while also having a well written profile blurb, as it cannot contain any obscenities and it must include a clear description that precisely describes the type of person a user is looking for and the gender must be specified.

With all of these amenities in place for both Tinder, match.com, and every platform that is designed for hookups or creating potential relationships, the downside to all of this is the profit that each platform wants to take in. As discussed before for Tinder, in order to access certain features that the application provides, there must be a payment. Match.com is not an exception to this. There are packaged deals that a user can subscribe to. Each package contains the same amenities but they differ in the duration of time you're allowed to access these services. Each package contains email read notifications, first impressions, and highlighted profiles. Members who have purchased a package are now subscribers while those who have not paid are simply members. Members are only allowed to search for matches, create a profile and appear in search results, post photos, send winks, have general mobile access, and receive text alerts for winks and emails. The features included for subscribers allow much more access for the user and gives a more in depth experience of the online dating experience.

The site exclaims that those who purchase the package for six months are guaranteed a match. There are also differences in the type of package a user will get. Every package stated above is in the bundle package plan. There are also more plans to view other than the bundled packaged plans such as the basic package plans. The basic plans are cheaper starting with prices as low as $19.99 a month for three months, $16.99 a month for six months (with a guaranteed match attached to it), and $15.99 a month for a year but they are not the ideal choices compared to the bundled packages. Users are unable highlight their profiles, decide on first impressions and be able to send

email read notifications. This tactic can be used to entice their customers by drawing in on the differences to make them more willing to pay.

Their reasoning behind making users pay for online dating is because every member is investing in themselves and their love. Their philosophy is that if there are users who pay for an online dating service then they have acknowledged that are searching for like-minded people and they are ready to monetarily invest into their love life. Match.com has listed five of the main reason to join their network; investment, safety and security, psychology, relationships and coaching, dedicated customer care, and the opportunity to attend exclusive single events. The claim for safety and security is that by communicating with people online and sharing personal information about yourself, this can create concerns that may arise due to a perception of lack of security. Anyone can chat as open and freely as they like.

One fault of match.com is the style of the layout. The layout of this site is aesthetically subpar. As discussed in chapter 1, a key strategy to creating apps or social networks is by appealing to your users with an aesthetic layout that draws the user in. A social network that is visually unpleasant raises the possibility of losing users. Social networks like Facebook, Instagram or Tinder have all achieved a certain aesthetic for their platforms, which are respectively different, but they each created something that satisfies a certain aspect to what their users want. Therefore, will all the are blurry ads all over and pixelated text that take over the home page and every other page, there is a lot that is going on with the layout and there is no clear cut way to navigate any page.

Match.com also has sister sites available to their users. chemistry.com and matchaffinity.com. These sites are similarly designed to one another but there is no

distinction that separates any of them. Each of their reasons to join their network does not highlight what they will specifically do for singles. Providing more opportunities for single people to meet each other and date one another will produce better opportunities to bring about matches, but the only downside about this is that potential users will have to pay each of these sister sites separately.

There are many similarities between match.com and the application that is being created for this project. Both are using the answers generated from each user to come up with a network full of people that will likely create a match. Finding the commonalities between people displays that they are more concerned with finding someone who has common interests similar to their own. Doing so would, in theory, incentivize a person to further explore a possible relationship since some of the similarities have been established.

The major difference that is most important to the users is that once a profile has been created, all of the amenities will be accessible to them, free of charge. Both Tinder and match.com offer packages for a certain price, which begs the question, are these social networks designed to find a partner for you or are the creators looking for a profit? Tinder is less meaningful in this than match.com because the users who match up are usually on this social network for a quick hook up. There is no real basis to match people up. Although match.com asks for answers based off of the certain aesthetic a user is looking for in their partner, which gives it a certain superficial feel where looks matter more than personality, the app being creating is not focused on the looks of a person that a user is searching for but more so creating dates.

Tinder and match.com are useful social networks to draw inspiration from. There are many similarities and differences between every platform but the differences that are highlighted within this chapter helps differentiate these social networks. There is no denying that the user will be able to recognize the the correlations but if it offers a smoother way of navigation, a stunning display, and incentives to date, then a user will be more galvanized to stay with the platform.

3
Stable Matching Algorithms

3.1 Stable Matching Problem

The creation of the stable matching problem came about through David Gale and Lloyd Shapely. The two contemplated together about the possibility of creating a self-enforcing job recruiting system that eliminated the self-interest of the companies (employers) and the students (applicants). Each applicant has an ordering on the companies they would like to apply to and each employer, once the applications start to arrive, forms a preference ordering on its applicants. Based on these preferences, companies will extend offers to some of their applicants, applicants choose which of their offers to accept, and everyone is off to their internship for the summer.

To think more informally about this, Gale and Shapely imagined a situation where college juniors were applying to internships for the summer. The problem that came up was if student A was hired by employer A then that student would accept the offer and their summer would be the start of a new beginning. But what if student A was offered another position by employer B - whose laid back, anything can happen attitude was preferable - and that student chose instead to work for that company? Student A would retract their acceptance from employer A and accepted an offer from employer B and now employer A would have to extend an offer to the applicants who were wait listed. All this, which is based on preference and self interest, might cause a spiraling system where applicants are retracting their acceptances and employers now looking for others to replace them in the hopes that they will accept and not alter their decision.

A situation can also occur just as bad on the other side. Suppose student B heard the story that student A told them and now calls up employer B and says “you know, I’d rather work for you than at employer C”. Seeing that this is very reasonable, employer B looks over student B’s resume and company B decides to accept student B. However, upon looking over the students resume, the employer comes to the conclusion that the student they already hired, student C, is less qualified than student B and prefer student B over student C. Seeing that employer B is a questionable company, they will find ways to retract their offers and offer it to other students.

Since situations like this can result in chaos, the reason for this occurrence is that it is not self-enforcing. Many people, both applicants and employers, can end up unhappy with the process just as much as the outcome. If people are allowed to act in their self-interest, then the possibility of everything breaking down is high.

There is, however, a more stable situation where self-interest works in on itself and prevents applicants and employers from retracting or redirecting their answers. If we consider another situation where employer D, who has already hired all of their applicants for the summer, receives a call from applicant D requesting that they accept them as an intern. Based on this case and all the offers already accepted, they can reply “Nope. Turns out we like all our applicants compared to you so there’s nothing we can do”. Or vice versa when a company is earnestly searching for their top applicants to accept and they all reply in “No, I’m happy where I am”, this creates a case where all matchings are stable. There are no other deals that can be made between applicants and employers.

So this was the situation that both Gale and Shapely came up with for this question: Given a set of preferences among employers and applicants, can we assign applicants to employers so that for every employer *E,* and every applicant *A* who is not scheduled to work for *E,* at least one of the following two things is the case?

(i) *E* prefers everyone of its accepted applicants to *A*; or

(ii) A prefers her current situation over working for employer *E*.

If one of the two cases presented holds, then the outcomes are stable: individual self-interest will prevent any applicant/employer deal from being made from behind the scenes.[14]

It is important to add that this is not the origin of this problem. Without knowledge, the National Resident Matching Program has a very similar procedure to the one Gale and Shapely presented. Using the same underlying motive, the program used

---

[14] Eva Tardos, Jon Kleinberg, “Introduction: Some Representative Problems” in *Algorithm Design* (Boston: Pearson Education, 2006).

an algorithm to match patients to hospitals. This system is still in use today with little to no change.[15]

An example of a more in-depth version of this problem is to simplify it to two individuals: male and female. Let's note that Gale and Shapely considered a same-sex Stable Matching Problem with a single gender. This is motivation by related applications, but is fairly different on the technical level. Given the applicant-employer application we're considering here, Gale and Shapely decided it would be best to focus on the version with two genders.[16]

We consider a set $M = \{m_1,\ldots,m_n\}$ of n men, and a set $W = \{w_1,\ldots,w_n\}$ of w women. Let $M \times W$ denote the set of all possible pairs of the form (m,w), where $m \varepsilon M$ and $w \varepsilon W$. A *matching* S is a set of ordered pairs, each from $M \times W$, with the property that each member of M and each member of W appears in at most one pair in S. A *perfect matching* S' is a matching with the property that each member member of M and each member of W appears in *exactly* one pair in S'.[17] In the situation that is occurring presently, a perfect matching corresponds to a way to have every man and every woman married to one another, in such a way that no one is married to more than one person. Single-hood and polygamy (or polyamorous relationships) have been ruled out in this case.

Once we have a established a set for both men and women, we now take into consideration each members' preferences. Every man $m \varepsilon M$ will now rank all women; it

---

[15] Eva Tardos, *Algorithm Design, 2.*

[16] Eva Tardos, *Algorithm Design, 4.*

[17] Eva Tardos, *Algorithm Design, 4-5.*

will be stated that *m* prefers *w* to *w'* if *m* ranks *w* higher than *w'*. This ordered listing is a preference list for *m.* There will be no exceptions for ties and the ranking system will be the same for women as it is for men.

Let's assume that there are perfect matches in *S*, is there anything that can go wrong? There is a situation that would be a cause for concern: There are two pairs of matches (*m,w*) and (*m',w'*) in *S* with our initial property that *m* prefers *w'* to *w* and *w'* prefers *m* to *m'*. If this is the case, then a situation can occur in which *m* can abandon his current marriage proposal and run off with *w'*. The set of marriages are not self-enforcing so we'll state that (*m, w')* is an instability pair with respect to *S* such that (*m,w'*) do not belong to *S* but each of *m and w'* prefers the other to their partner in *S.*[18]

The goal is to then list a set of marriages with no instabilities. Let's state that a matching *S* is stable if (i) it is perfect, and (ii) with respect to *S,* there are no instabilities. Such a statement brings about two questions:

- Does there exist a stable matching for every set of preference lists?
- Given a set of preference lists, can we efficiently construct a stable matching if there is one?

For example, let's illustrate these definitions. We will consider two of the following instances for the Stable Matching Algorithm. First, suppose we have a situation where we have two sets of men {*m,m'}* and {*w,w'}.* Then the preference lists are as follows:

- *m* prefers *w* to *w'*
- *m'* prefers *w* to *w'*
- *w* prefers *m to m'*

---

[18] Eva Tardos, *Algorithm Design, 5.*

- *w'* prefers *m* to *m'*

Let's intuitively analyze this set of pairs. This set of preferences represents complete agreement on both the men and women side. The men agree on the order of the women and the women agree on the order of the men. With the sets of matches given here, there is a unique matching which consists of the pairs (*m,w*) and (*m',w'*). The reason for this is because (*m,w*) would form an instability with respect to this matching. The perfect matching, however, will consist of (*m',w*) and (*m,w'*) while these are not stable matchings. This means that *m* and *w* want to leave their respective partners and pair up together.

Now if the pairs are a bit more intricate, we will see a different preference order

- *m* prefers *w* to *w'*
- *m'* prefers *w'* to *w*
- *w* prefers *m'* to *m*
- *w'* prefers *m* to *m'*

The men preferences are meshing with the other in a perfect way as they both rank the women differently. This is the same for the women but although they match up separately, there are conflicts between the men and women's ranking as they clash.

We currently have two stable matchings. The two pairs (*m,w*) and (*m',w'*) are stable because both men are as happy as possible with their given partner. When this is the case then neither partners would break a part from their matched partner. There is also another instance when (*m,w'*) and (*m',w*) where they're both a stable matching for the complementary reason that both women are happy as they can be.

3.2 *Designing the Algorithm*

It is important to show that there exists a stable matching for every ordered lists of preferences between the men and the women. Let us show that there is an efficient algorithm that answers the second question stated previously. If we can consider the basic ideas that would give the algorithm a motive: Initially, both the men and women are unmarried. If an unmarried man *m* wants to marry the woman *w*, who also places first on his preferences, he proposes. Can we say that the matching pair (*m,w*) is in the final stable matching? Not necessarily because a man *m'* whom *w* could possibly propose to her in the near future. But it would not be wise to immediately leave *m* because *w* has not received any other proposals from another man *m'* who is ranked as high as *m*. The best decision to make for *w* is to keep *m* in a intermediate state such as an engagement. Here is an example of how the Stable Matching Algorithm works:

Initially all $m \varepsilon M$ and $w \varepsilon W$ are free
While there is a man *m* who hasn't proposed to every woman
    Choose such a man *m*
    Let *w* be the highest-ranked woman in *m's* preference list
        to whom *m* has not yet proposed to
    If *w* is free then
        (*m,w*) become engaged
    Else *w* is currently engaged to *m'*
        If *w* prefers *m'* to *m* then
            m remains free
        Else *w* prefers *m* to *m'*
            (*m,w*) become engaged
            m' becomes free
        Endif
    Endif
Endwhile
Return the set *S* of engaged pairs

Although the algorithm is very simple to state, it is not obvious that a stable match or even a perfect match will be returned.[19]

3.3 *Stable Roommates Problem*

The Stable Roommates Problem draws a lot of similarities to the Stable Marriage Problem since it stems from the Gale-Shapely algorithm. The roommates problem is essentially the stable matching problem, however, the difference is that it uses only one set.The algorithm breaks down naturally into three phases. The first phase of the algorithm is based on a sequence of proposals that are made to one another. This sequence of proposals is followed by each individual following their strategies of obtaining their preferences:

(i) If Dave receives a proposal from Mary then

(a) he will reject it at once if he already holds a better proposal. For example, if he receives a proposal from a higher ranked person or;

(b) he will hold it for consideration while also rejecting those who are ranked lower than the current person.

(ii) Dave proposes to the others in which they appear in his preference list, stopping when a promise of a proposal has been made; any rejections would therefore cause Dave to continue immediately his sequence of proposals.

The individuals will begin their proposals sequence one at a time and with each person's first proposal will cause a chain reaction of acceptances or rejections with the

---

[19] Eva Tardos, *Algorithm Design, 6.*

hopes that it will eventually terminate only when someone receives their first proposal or after successfully matching multiple pairs with one another.[20]

Let's consider an example where there are different people looking to room with one another based off of their preferences.

1 4 2 6 5 3

2 6 3 5 1 4

3 4 5 1 6 2

4 2 6 5 1 3

5 4 2 3 6 1

6 5 1 4 2 3

According to the rules that were previously stated, the following will therefore take place:

1 proposes to 4; 4 holds 1;

2 proposes to 6; 6 holds 2;

3 proposes to 4; 4 rejects 3;

3 proposes to 5; 5 holds 3;

4 proposes to 2; 2 holds 4;

5 proposes to 4; 4 holds 5 and rejects 1;

1 proposes to 6; 6 holds 1 and rejects 2;

2 proposes to 3; 3 holds 2;

6 proposes to 5; 5 rejects 6;

---

[20] Robert W. Irving, *An Efficient Algorithm for the "Stable Roommates" Problem (United Kingdom: Salford, 1984) 577-580*

6 proposes to 1; 1 holds 6;

This phase of the algorithm will terminate either

(i) with every person holding a proposal or

(ii) with one person rejected by everyone (as would happen if the matrix contained 4 different elements).

In case (ii) you can see that everyone now has a proposal except for the person rejected because everyone in the matrix has now rejected them. Since this is the case then everyone else has made a proposal.

The second phase of the algorithm mainly involves the handling of the reduced preference lists. As in the example above, some lists are able to contain more than one person. The second phase involves further reduction of these preference lists and can be repeated several times. It will only terminate once the one of the following two terminating rules have been recognized. Either one person runs out of people to propose to, thus causing the lists to have an instability, or all the preferences shrink so that every person within the matrix now has a match in which a stable matching can occur. After process of elimination is complete, the third phase of the algorithm consists of finding a stable match for each participant in the given set.

We first find a participant who has more than one choice. We take note of their second preference, X, and their last preference, Y, of X. We will repeat the previous step until the starting participant appears again. Finally, we look at every second preference and last preference and reject symmetrically. The following algorithm will illustrate how all three phases work:

```
While there are unmatched people
        Let i = the smallest value such that a_i is unmatched
        a_i proposes to his favorite roommate a_j who has not rejected
        him previously
        if a_j has not received a proposal before then
                a_j accepts a_i
        else
                if a_j prefers a_i over his current match a_k then
                        a_j accepts a_i
                        a_k reject a_j
                else
                        a_j rejects a_i
                endif
        endif
endwhile
for all accepted proposals
        reject all (a_j,a_k) where a_j prefers a_i over a_k
end
for all cycles (p_1,…,p_n) and associated second preferences (q_1,…,q_n)
such that:
        q_i is the second preference of p_i
        p_i +1 is the last preference of q_i
        p_n ε {p_1,…,p_n-1}
                for i = 1,..,n-1
                        q_i rejects p_n+1
                end
end
```

With the introduction of the Stable Matching algorithm, both marriage and roommate, it is important to state why these algorithms are essential to this project. Its uses are vital to how users are matched up because it doesn't solely lie on the questionnaire to match two individuals or a cluster of individuals to select from. By utilizing these algorithms, we can see how matches can occur within a network.

Let us note that the matching of two users is not the final marker for those two users, as this is still a social network. The idea of marriages and proposals do not

necessarily pertain to this structure but it important to see the other ways in which matching can take place.

4
Hybrid Algorithms

4.1 Algorithm Functions

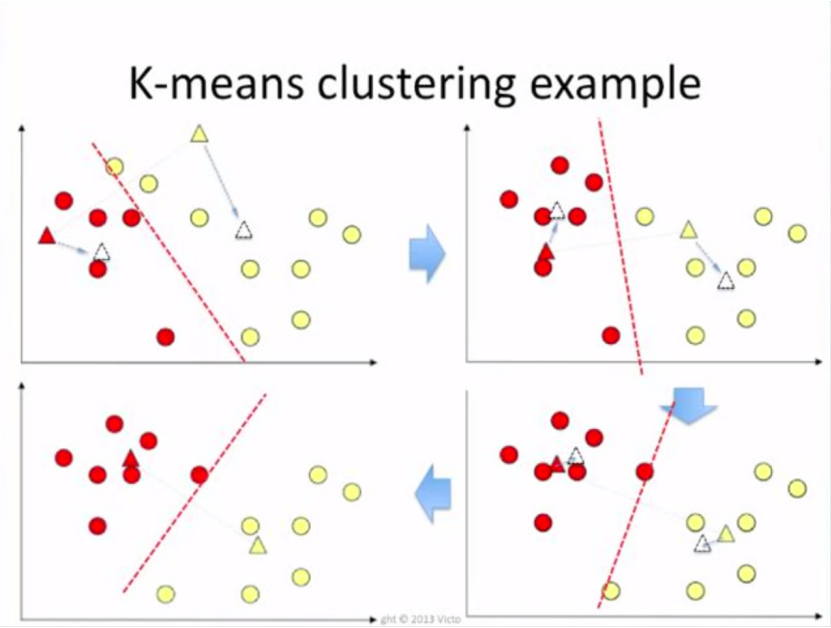
The algorithms presented in the previous chapter displays the ways in which matchings can be made. The Stable Matching algorithms will be specified to highlight how each of them will come into affect to created the best possible matches. The modification of the algorithm will take answers from the questionnaire to generate what each user has in common with one another, but the matchings that occur are suboptimal matches. If there are multiple users who have more in common with a member than those that have less in common, then it would be ideal to place those who rank higher in the forefront. Because a users preferences are searching for the best possible choice rather than the ones who are least compatible, an algorithm will be needed to rank the users to determine the best clusters of user.

Before any of the algorithms are implemented, it is important to cluster a users network. Given a set of random number of individuals who are unmatched in a network, there will be an additional part to the algorithm that will cluster everyone based on the preferences given from each user. The first step to the process is clustering all the users together in the appropriate sets. There are many ways to cluster points of data on a given plane, such as finding the euclidean distance, k-means clustering, the edit distance, the Jaccard index, otherwise known as Jacquard similarity coefficients, and the cosine distance formula.

For this project, it would be ideal to use k-means as a way of clustering data points. The data points in this case will be the men and women who have access to the application. In a given set, *k-means clustering* will have *n* data points in d-dimensional space R^d and an integer *k*. The problem then is to determine a set of k points in R^d, called centers, so as to minimize the mean squared distance from each data point to its nearest center. For k-means clustering, we can use Lloyd's algorithm which is an algorithm that is easy to implement. First, we select *k* arbitrary points as the centers. After the centers have been selected, we then find the most reasonable assignment of data points to the center. The best way to assign data points to centers is to assign each data center to the closest center. This is known as the centers-to-clusters step of the algorithm. After the clusters have been constructed, the centers must be in a position that is closest to the the clusters. The process of moving the center points to their proper position within the clusters is known as the cluster-to-center step of the algorithm.[21] When new data points arise on the graph, we will keep assigning those

[21] https://www.youtube.com/watch?v=9rp1pzYn3hY

data points to its nearest center and then move the centers to the center of the clusters. The algorithm will terminate once all points have stopped moving but for the application, there will always be new points added the clusters. An example of this can be seen in figure 4.1.



4.1 An example of how k-means clustering separates data points

The way in which the clusters will be differentiated is when users establish their genders in the questionnaire. Once the gender is established, both male and females will be represented on the plane as two separate data sets. Those who identify as male will be separated from those who identify as female. The k-means algorithm will then run to create the clusters for both men and women.

Once each cluster has been created, we now have to separate the current clusters, which seems a bit challenging but it can be easily handled if one of the following occurs: separate each clusters by taking the average of all users and grouping them together based on the highest average to the lowest averages from the answers

given in the questionnaire. We could also find the median number of preferences to determine the most suitable match, which will then place the potential matches in consecutive order; we can also potentially use mid-range, also known as mid-extreme, which will take the average of the lowest ranking person with the average from the highest ranking person and find the sum of those two integers. Once the sum is calculated, the numerator is now divided by 2 to detect the best possible match, and finally, the mode. The mode finds the most populated numbers that appear in a given set of numbers.

Although these ways of finding the best possible match for a user is very simple, its uses can drastically simplify the process of discovering a match. Therefore, what would be the most simple and efficient way of discovering a match would be to find the average of all the potential matches. The reason for this is because when the clusters have been made, calculating the averages and choosing the best possible match can drastically limit the options available for a user. Once the averages are calculated, we can now view each cluster with the highest average down to the lowest.

Once the averages are detected, here is where the stable matching algorithm comes in to minimize all the users and detect the most suitable match. After the averages are calculated, the highest averages of every cluster is then taken out of the main cluster and then paired with the member who ranks closely to them. Since the Stable Marriage algorithm contains two sets with multiple elements within each of them, for this algorithm, the two sets will remain but there will be one set with one individual user and the other set containing the users with the highest averages. After the sets have been established, the algorithm will now find the user with the most

commonalities. It is important to note that the users that are generated from these clusters do not necessarily prefer the user that they are presented to. The Stable Matching algorithms are formally stated such that for any two people who have been paired together, there must be no other possible matching that both people would rather be in. However, as these matchings are not permanent in a dating pool, it will only create a temporary match for a date.

The Stable Roommate Algorithm is used in a different yet similar fashion as the Marriage Algorithm. Since it only involves one set, the idea of it being used to match men and women together would not be efficient enough to produce a stable matching. With this idea in mind, it would be most efficient to use the algorithm for same sex matchings. By way of illustration, users can be interested to their same sex and/or the opposite sex. For those that do have those specific preferences in their partners, the algorithm will run on each of the individual gender cluster to generate the ‘best roommates possible’. The k-means algorithm will create a cluster within the existing cluster. With this being said, matchings can transpire with both same-sex interests and opposite-sex interest. If a user has chosen to be interested in both, the Stable Marriage Algorithm will only come into affect when presenting someone of the opposite sex while the Roommate Algorithm presents someone of the same sex.

The algorithms will not the be the main source for determining a match. A user will not be forced to have this self-enforcer algorithm create a match. All of this takes place in the backend database. It will be presented as an option to view a recommended match for a user. If a user would rather explore their network to get the full experience then the option is always available to them.

4.2 *Database Analysis*

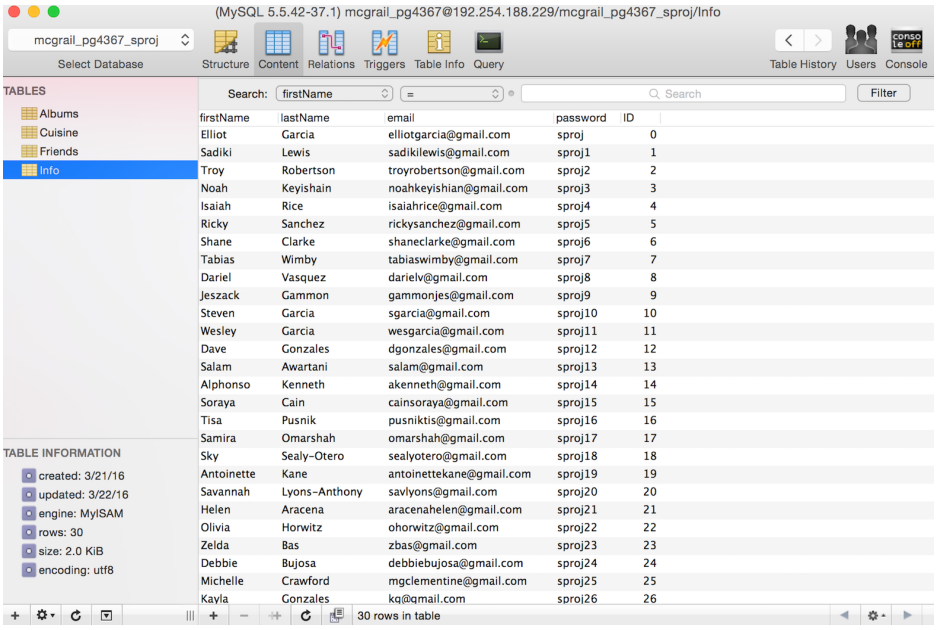
The information that is needed for all of this to occur is the data that has been stored into the database. What the database stores is a minimal amount of information required from every user. The information that we are mainly examining is seeing how matches can occur, either through the self-enforcing algorithm or not, and more importantly, what caused each user to become a match. Analyzing what created a match is important to see which of the categories was the cause that brought the two users together.

The data can also reveal the problems that exist when dating pools are comprised of solely men and women. For example, if there are 15 women standing alone in dating service app with around 300 men, we could see how the algorithms in place could easily handle a situation that is being presented. The algorithm will take the 15 women and cluster them in to one set and the 300 men are clustered into another set. Then it will divide the men up accordingly based on the ratio of women to men and the two cases are as follows:

> Case (i) there will now be 20 groups of men available for the women to choose and decide which man within the pool they find most desirable.
> Case (ii) after the groups have been divided up, the averages will be taken from each group to determine which group is more suitable for each woman. After collecting the averages, the Stable Marriage algorithm will run on each group to decide the best match for each woman.

This can help ease the problem of a dating pool with an existing imbalance. Although there many more men who are left without a date, at least there are 15 men and women who will leave with one. The purpose is not to create a permanent match but rather create suboptimal matches for a date. The imbalance will always exist as this is a dating pool but because there are no permanent matchings that are being made, many dates can happen if the original suboptimal match didn't work out.

The information that is stored into the database is crucial for this project. As stated before, when singing up for the dating service, there are numerous questions being posed to the user to get a sense of who they are. Gender, interests in food and other specific qualities are used to illuminate a certain aspect of their character to others. This information is then manipulated by the algorithms to find a stable match for a date. An example of the information stored is shown by figure 4.2.



firstName	lastName	email	password	ID
Elliot	Garcia	elliotgarcia@gmail.com	sproj	0
Sadiki	Lewis	sadikilewis@gmail.com	sproj1	1
Troy	Robertson	troyrobertson@gmail.com	sproj2	2
Noah	Keyishain	noahkeyishian@gmail.com	sproj3	3
Isaiah	Rice	isaiahrice@gmail.com	sproj4	4
Ricky	Sanchez	rickysanchez@gmail.com	sproj5	5
Shane	Clarke	shaneclarke@gmail.com	sproj6	6
Tabias	Wimby	tabiaswimby@gmail.com	sproj7	7
Dariel	Vasquez	darielv@gmail.com	sproj8	8
Jeszack	Gammon	gammonjes@gmail.com	sproj9	9
Steven	Garcia	sgarcia@gmail.com	sproj10	10
Wesley	Garcia	wesgarcia@gmail.com	sproj11	11
Dave	Gonzales	dgonzales@gmail.com	sproj12	12
Salam	Awartani	salam@gmail.com	sproj13	13
Alphonso	Kenneth	akenneth@gmail.com	sproj14	14
Soraya	Cain	cainsoraya@gmail.com	sproj15	15
Tisa	Pusnik	pusniktis@gmail.com	sproj16	16
Samira	Omarshah	omarshah@gmail.com	sproj17	17
Sky	Sealy-Otero	sealyotero@gmail.com	sproj18	18
Antoinette	Kane	antoinettekane@gmail.com	sproj19	19
Savannah	Lyons-Anthony	savlyons@gmail.com	sproj20	20
Helen	Aracena	aracenahelen@gmail.com	sproj21	21
Olivia	Horwitz	ohorwitz@gmail.com	sproj22	22
Zelda	Bas	zbas@gmail.com	sproj23	23
Debbie	Bujosa	debbiebujosa@gmail.com	sproj24	24
Michelle	Crawford	mgclementine@gmail.com	sproj25	25
Kayla	Gonzales	kg@gmail.com	sproj26	26

Figure 4.2 An example database of mock data.

It is important to note that the data collected is mock data. Here we see the general data that is collected from signing up. An email is needed to confirm that the users exists and every user will have an ID number associated to them. This is important for linking a user to all the other information because it corresponds to all the information in the other tables that carry ID numbers. The gender column will establish the sets that will help run the algorithms. By gendering people, it is easier to place them into the two sets so that the marriage algorithm can run its course to create a match.

Based on the general information given from every user, they are each assigned an ID number. The cuisine table in the database is the table in which the averages are taken from. As seen in figure 4.3, every user entered a range of numbers for their

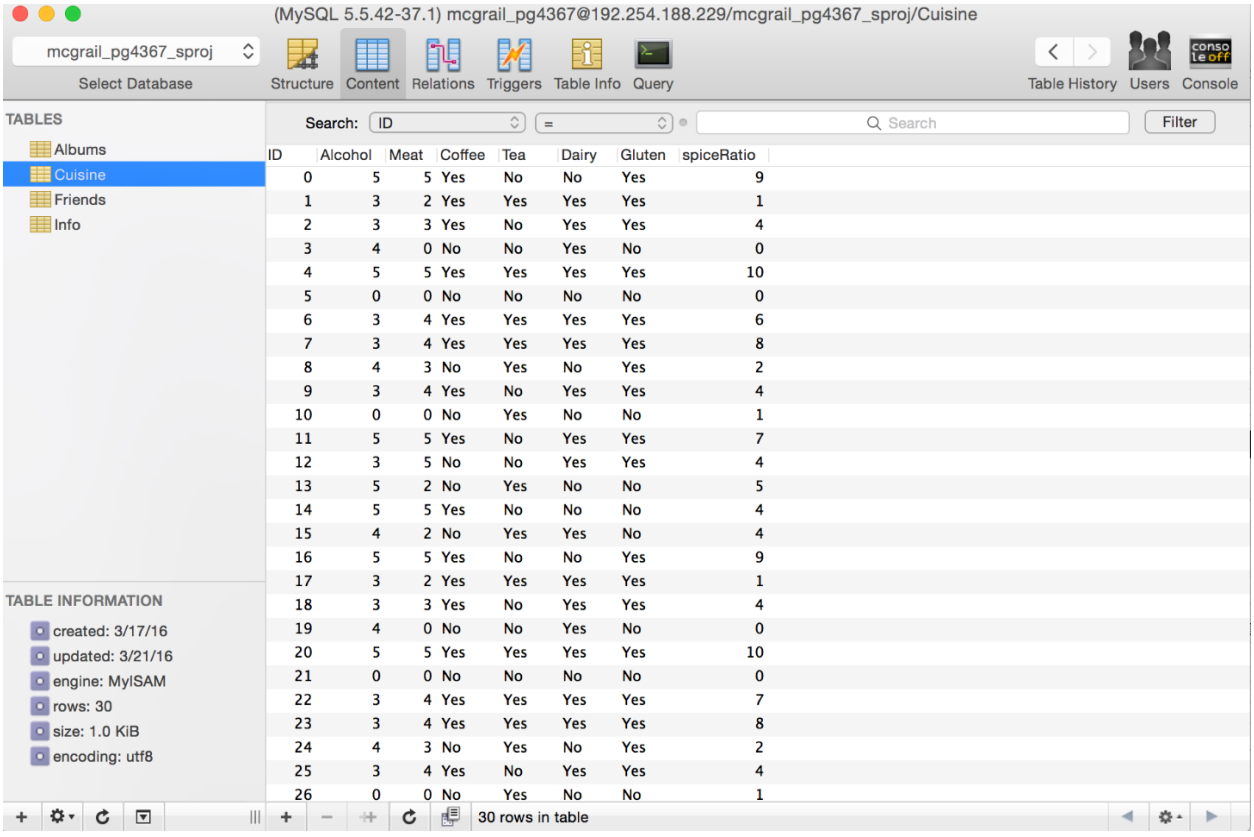


Figure 4.3 An example table for the cuisine table

preferences in food. It will be those numbers that the averages will be taken from to find the best average for a date.

The other tables within the database are used as a way to keep track of every user. It is important to see how active a user is by observing if a user is utilizing their photo album and how many friends are within their network. When viewing the activity of a user, we can see their personal engagement with the application. In figure 4.4 we see that the friends table is empty. The reason for this is because the users have not yet begun any activity and thus, no networks were created. Once again, the ID number will correspond to whichever user is associated to that number so if they were added or blocked, the ID number will indicate who it was. The common interests column marks the similarities between each user. These similarities are the answers that match closely up with a user who is within their network. This table is an example table for a single user. Figure 4.5 displays what a friends table would display.

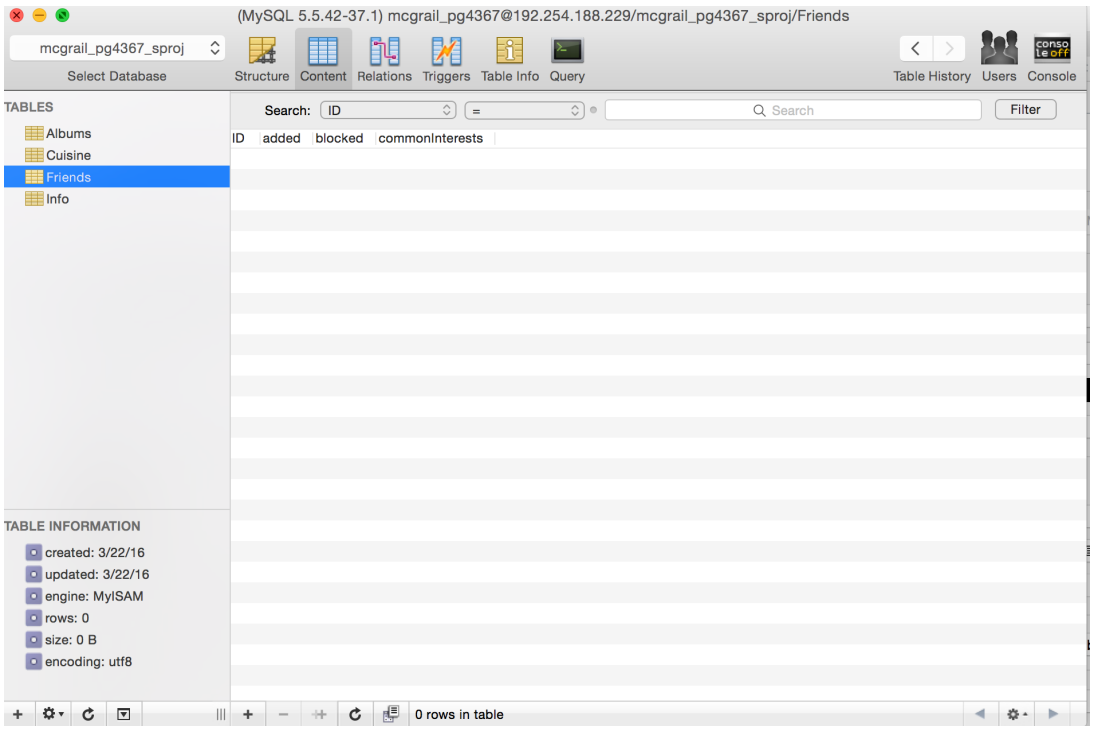


Figure 4.4 Here is an example of a table containing the friends of a user

Every user has the option to upload a variety of photos to one album to display themselves. The photos themselves are not entered into the database but the album names are, even when the names of the albums have been changed. The importance of keeping track of the photo albums are once again to see how engaged the users are with the application. If they choose to upload photos, it raises the prospect for a user to find a match. Figure 4.5 shows the database containing the photo album names and the ID number. For this table in particular, there is no information that can be gathered. The information that would pertain to the photos

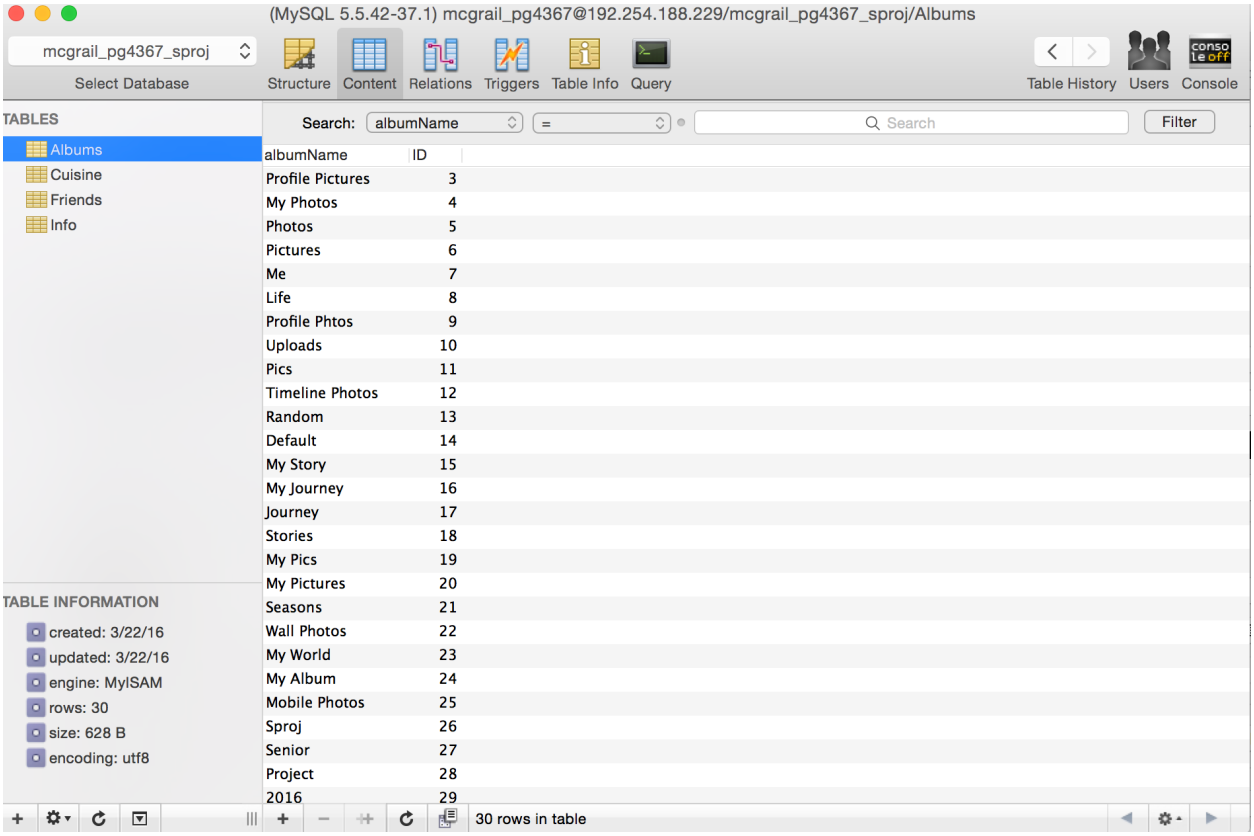


Figure 4.5 of the mock photo album name table.

would be of no importance to us as we are only considering this table to view the activity of the users.

Once all the information has been gathered, it is important to note what will be done with this information. As discussed in chapter 1, many social networks such as Facebook for example, collect information from their users based off of what is being talked about through memes or trending topics or how people feel about the social climate from their networking platform. There are many ways in which people can express their agency to gain a greater sense of self. For example, Facebook conducted an experiment on their users by analyzing their moods and displaying statuses, images, and advertisements that pertained to their mood, which was established through what the users were posting on their timeline.

There are many things that the information from the database can tell us but what we are interested in is the likelihood of someone being matched based off of their similarities or if this does not appeal to the users, we will simply analyze the locations that are more popular. What we can also gather from the users are which cuisines from the table are most likely to create a match. Once those matches have been created, we would like to see which locations are more popular than others. Since each user has agreed to share their location, we are able to see which cities, towns, and or villages are more likely to have a successful match. We will be able to see which foods bring people together internationally and regionally. Also, we will see how effective the Stable Matching Algorithms are for matching people and the odds of people accepting the offers presented to them. All this information gathered can lead to what people are actually looking for when they are out on a date.

5
Conclusion

The purpose of this project was to explore the best possible way to match two people up under the scope of food and socialization. We presented and analyzed two algorithms for creating matches such as the Stable Marriage and Roommate algorithms. These algorithms exhibited to match any two individuals within a dating pool regardless of gender or in the case of this project, cuisine preferences. A second critical component of this project was to create a hybrid algorithm from existing algorithms to create the best possible match for a date within a dating pool. Many things were taken into account - gender, cuisine preferences, location, etc. - to create the best possible dating pool for a user. Similar social dating networks were compared to see how the existing competition creates matches for their users. The in-depth discussion about the algorithms used for this project presented a different way to take opposite sex and same sex people within the same dating pool and match them together. All of this wrapped together created this project and hopefully it is taken a step further in the future.

Chapter 1 introduces the purpose of the project and analyzes how cuisine can bring people together. Mealtime, which is a critical cultural component, can be regarded as a cultural site where people learn how to socialize no matter the location. These sites are used to shape people for interaction with one another and we learn the ways to socialize at the dinner table, whether it is taking place at the home table or at a restaurant. When numerous participants are engaging in these activities, this creates a cultural site, and in this case, it is for mealtime. In these sites, there is an expectation on how people should act with others. Every site has a certain social expectation from their participants which is the reason why people behave in a certain manner when they're gathered together at these sites. We looked at how Twitter and emails can gather information from their users. The chapter then transitions into explaining how food and socialization can be transmuted into an application that uses these cultural sites as a way to give incentive to bring people together. The mock screens of the application representing the ways in which people can meet through such an application.

The second chapter explores the existing competition that uses location and questionnaires to match people together. Here, we looked at Tinder and match.com as our main source of competition. Although there are many more dating sites, these two were used because of the similarities in how people are matched and the popularity of these platforms. Both networks were highlighted for their similarities while also drawing on the differences in how people are matched together.

The third chapter discussed the algorithms that will be used in order to produce a match within an existing dating pool. The Stable Matching Algorithms, both marriage and roommate, are used as a self-enforcing algorithm that eliminates the users self-

interest in choosing a date. After creating clusters using K-means, the Stable Marriage Algorithm is used on two sets, in this case the sets are contained with men and women, and the averages are calculated to determine the best suitable matches for a given user. The Stable Roommates algorithm will operate on each set individually to determine matches for same sex couples. For those who are interested in users who are of the same gender, there will be a cluster created within the clusters and the averages will be taken from those clusters, followed by running the Stable Roommates algorithm to create the best pairs within a set. The fourth chapter then explains how the hybrid algorithm will come into affect for the project. It later looks at how the users information are stored into the database and how the data is manipulated in order to produce a match.

The importance of this project is to reveal various ways in which people can interact with each other and exploring various ways to match people together. This project started as a means to see how people can connect through cuisine. With the idea of a mobile application to get it started, it evolved into examining ways in which people interact with one another under the context of cuisine preferences. What this project can lead to is also of great importance because it shows the greater potential of finding a match within a dating pool. The chances of finding a date is much higher than waiting to find someone who peaks your interest. Although a permanent match is not guaranteed, a date is still likely to occur. Also, what this project can also show is which major cities are using the application. By tracing the users locations, we will be able to view which locations are more populated than others.

If there was more time allotted for this project, the results would have been more striking. The data used in the database would have been entered with real people giving their honest opinions. Because of legality issues pertaining to peoples personal privacy and safety, I was unable to gather information from them because of time. If this was not an issue, the results taken from everyone would have revealed actual preferences to match people together, which is something that the mock data did not. Furthermore, if the application was complete, then the app would have been available for campus use. This would have been important to get real life answers from students on Bards campus. Although the dating aspect would not have been of concerned for their purposes, it would have been interesting to see who was matched together and for what reasons. The launch of the application would have also showed if the Stable Matching Algorithms would have been affective in matching people together to their likings. The rate of acceptances or rejections would have been of particular interest because it would have determined how effective the algorithms were.

If there was a certain aspect to focus on in this project, it would have been the development of the application. Because of time constraints and other restrictions, the application was not able to be fully developed. App development is a time consuming process which requires a team of programmers to accomplish over a period time. To complete a project single-handedly of this caliber within 10 months is near impossible. What would have been more interesting is to see how people would have answered the questionnaire. The reason for this interest is because the variety in answers I would have received would have been remarkable. Being more socially inclined, the application also poses a beneficial tool to simply date those who you are interested in.

References

- Ferrara, Emilio, Mohsen JafariAsbagh, Onur Varol, Vared Qazvinian, Filippo Menczer, and Alessandro Flammini. *Clustering Memes in Social Media.* Accessed October 10, 2013. http://arxiv.org/abs/1310.2665.
- Irving, Robert W. *An Efficient Algorithm for the "Stable Roommate" Problem*. Salford: University of Salford, 1984.
- Kleinberg, Jon, and Éva Tardos. *Algorithm Design*. Boston: Pearson Education, 2006.
- Ochs, Elinor, and Merav Shohet. *The Cultural Structuring of Mealtime Socialization*. Los Angeles: Wiley InterScience, 2006.
- Russell, Matthew A. *Mining the Social Web*. 2nd ed. O'Reilly Media, 2013.