ARE ALL DATING APPS CREATED EQUAL: COMPARATIVE SIMULATIONS OF MOBILE DATING APPLICATIONS THROUGH LEVELS OF MULTIPLICITY

Fatma Abdulwahab
Collin Henley
Steven Hoang
Liam McKenna
Joseph A.E. Shaheen, Ph.D

Department of Computational and Data Sciences
George Mason University
4400 University Drive
Fairfax, VA, USA
jshaheen@gmu.edu

ABSTRACT

This experiment seeks to compare the effectiveness of simple vs. complex mobile dating application interfaces in acquiring matches for users. Simulations were created to replicate today's most popular location-based real time dating mobile applications (LBRTD) based on varying levels of behavioral complexity. Agents in each simulation (n=1000) are assigned gender and randomized attribute values. Male and female agents engage in pairing activity based on rule-sets drawn from academic literature. We observed a direct relationship between the complex behavior (multiplicity) allowed by simulated platforms and the 'Match' distribution as well as network structure. We validate and expand our findings with a robust statistical model.

Keywords: online dating, social networks, agent-based modeling, mobile applications.

1 INTRODUCTION

Location based real time dating (LBRTD) applications have seen an exponential increase in user adoption in recent years. With 15 of the most popular apps acquiring 247 million downloads in 2018 alone (Koch 2019), LBRTD has become one of the most popular platforms among singles to meet other singles and develop relationships. This domain has also proven to be a high revenue generator for businesses garnering over $2.1 billion in gross sales and projected to reach $3 billion in the coming years (Sumter, Vandenbosch, and Ligtenberg 2017).

This growth is driven by several factors, one of which is changing social norms and a common acceptance of online dating. In a study conducted by Pew Research Center, 44% of Americans indicated that "online dating is a good way to meet people" in 2005, and 29% agreed that "people who use online dating sites are desperate" (McGrath 2015). When participants were asked the same set of questions in 2013, this trend has already begun to change, yielding 59% and 21%, respectively, and consequently indicating that "online dating has lost much of its stigma".

Today's mainstream mobile dating applications offer common features and and a market-borne universal user interface, but continue to vary in levels of multiplicity–that is–the number of features available to users to achieve the same result, finding a suitable match.

One salient example is the simple interface offered by Tinder, a Match Group Inc. dating application that is credited with bringing LBRTD further into the mainstream. Tinder is estimated to have a current rate of 15 million matches per day (Tyson, Perta, Haddadi, and Seto 2016). Typically, after initial download, users are prompted to create a profile to showcase to potential matches. These profiles consist of images uploaded by the user and a single field for self-description. After users are satisfied with their 'profile', they can begin to choose preferences based on age, gender, distance, and location. The next step of the process involves specifying a criteria for preferences. Users can sort through potential profile stacks shown to them one at a time by the app's recommendation algorithm. Users then perform an action to either accept or reject the profile they have been shown (accept or reject). After each individual action has been performed, the user is shown another profile in the digital deck where the accept/reject process can be repeated. When two users both accept (like) each other through interface, a notification is sent to each user, and they are both able to communicate. This is deemed a 'match'.

As Tinder's popularity grew, other mobile dating apps emerged, each with their own rule-sets and interfaces. Another mobile application, OKCupid, gives users multiple ways to match extending all the functions of Tinder in addition to a stronger filter and search functions. Users also have the option to create extensive profiles that specify ethnicity, political views, body type, desire to have children, type of relationship desired, and other demographics. Prompts to give potential matches additional insight and act as conversation initiators can be posted on users' displayed profile. Users that are shown profiles adopting these prompts can send acceptance notifications of those prompts or post comments. Either action will result in a notification on the receiving user's profile–allowing for a direct line of communication where the receiving user can view the person accepting their profile without the need to sort through other profiles on the app interface. Additionally, digital profile decks that a user has already accepted are stored in a database that can be accessed by the user to allow for direct messaging attempts in order to assist the individual in gaining visibility. Users that desire an even more complex approach can use CoffeeMeetsBagel, another LBRTD application which has all of the features of Tinder and OKCupid with another added layer of multiplicity.

Given the wide array of apps, each with their different levels of multiplicity, analysis is necessary in determining whether more features generate more user matches. One of the primary key performance indicator of a successful dating app is the number of matches the app acquires for each user. Thus, we motivate our study of multiplicity of online dating applications–the many ways a match can be made.

In this paper, we construct two dating app simulations with a low multiplicity level representing apps such as Tinder, and another with a higher level of multiplicity representing applications such as OkCupid. We dub them Multiplicity Level 1 and Multiplicity Level 2. These simulations will be assigned rule-sets with the purpose of replicating their respective real-world counterparts. Our approach combines agent-based modeling and classical social network analysis as our go to analytical frameworks. Our constructed networks contain various edges/ties that are set to reflect acceptance actions and clicks from each individual user. We define a 'like' as a one sided acceptance of anothers' profile, a 'match' as a dyadic agreement, a 'dislike' as a rejection, and a 'message' as a directed tie. Through implementation of these simulations we investigate the contribution of various mechanisms, or lack of, in generating matches.

1.1 Behavioral Background

Homophily In Social Networks

The principle of homophily implies that "people's personal networks are homogeneous with regard to many socio-demographic, behavioral, and intrapersonal characteristics" (McPherson, Smith-Lovin, and Cook 2001). The social dynamic principle infers that people like those who are like themselves. This cognitive process has been found to be correlated with a sociological concept McPherson calls "constructuralism", the assumption that possibility of shared knowledge creates a higher likelihood that an interaction will take place. The assumption of shared knowledge would be associated with ease of communication, shared interests, and preferences that create a smoother interaction between individuals. In the context of dating, perceived similarity has been shown to follow a positive correlation between attraction and relationship satisfaction. A study conducted by (Fiore and Donath 2005) measures the effect of homophily by sampling a pool of users on an unnamed online dating application and analyzing their preferences for potential partners similar to their own preferences. Similarities like desire for children, level of education, and physical appearance were found to account for a better likelihood of a match than random. We rely on this sociological framework and incorporate it in our simulations, though admittedly, other sociological processes are likely at play. Still, homophily is a strong foundational driver for our virtual experiments. Specifically, we incorporate general notions of age, physical attractiveness, ethnicity and gender.

2 MATERIALS AND METHODS

2.1 Agent Attributes

Mating behavior varies based on the gender of the individual. Theoretical perspectives of evolutionary biology comprises the notion that the evolution of organisms occurs through differential gene: That is, obtaining an advantage over others of the same gender and/or optimizing survival of environmental conditions (Greenlees and Mcgrew 1994). Sexual selection consists of two processes: intrasexual selection and intersexual competition. Intersexual competition is defined as preferences members of one sex have for members of the other sex that possess certain valued characteristics (Greenlees and Mcgrew 1994). In both of our models, agents favor homophily as referenced in McPherson et al's (2001) study. Homophily between agents is determined by similarities between attractiveness and ethnicity. These similarities increase the probability that an edge will occur resulting in a higher likelihood of a match. The mechanics of ethnicity and physical attractiveness are subject to these effects in our models

However, these attributes contrast with patterns in age difference. "Men who select younger women should be reproductively advantaged over men who select older women" (Greenlees and Mcgrew 1994). Thus, male agents were given a preference to slightly younger female agents within a standard deviation of 3.31662, and female agents were given a preference to older males within a standard deviation of 2.90659 (Conway, Noë, Stulp, and Pollet 2015). We rely on data (Conway, Noë, Stulp, and Pollet 2015) and use a Gaussian distribution to give an agent relative value in their match power when matching against other agent. The further away the difference in age between two nodes is from the mean preferred age difference, the lower an agent will score in matching power, thus contributing to a lower likelihood of other interactions. For example, a female agent who is 4 years younger than a male agent will have a much better score with the male agent than if she were to be older than him.

Male agents were given a higher weight for attractiveness preference in accordance with Greenlees and McGrew's (1994) findings. Their study showed that "selection of women directly on the basis of age is problematic. Because of potential deception and inaccuracy, men should have evolved a preference for secondary correlated with reproductive value. Physical appearance is a strong cue to age and health, which in turn are indicative of reproductive value. Men who prefer physical traits that are good predictors of reproductive value should gain a selective advantage over their competitors" (Greenlees and Mcgrew 1994). Though education and profession are not included in our model, if included in future works, agents would be given preferences for education, and profession.

2.2 Multiple Regression Quadratic Assignment Procedure

The Multiple Regression Quadratic Assignment Procedure (MRQAP) is used to test for significance in an observed correlation (Krackardt 1987) where dependency between two or more dyadic relations may exist. It is a non-parametric, permutation-based test that preserves the integrity of the observed structures" (Krackardt 1987). This approach was used during an experiment conducted by Mantel (1966) to identify geographic clustering of diseases (Mantel 1967). MRQAP is particularly useful when calculating the permutations of network correlations. Since our samples, rules, attributes and mechanism are not identically and independently distributed we must rely on this powerful analytical tool-set. We omit a discussion of the nature of this method at this time, but direct the reader to aforementioned citations.

2.3 Model Mechanics

We use the Python module NetworkX to build our model and run our simulations. To issue a direct comparison agent attributes were kept to a minimum and consistent across both model types. We instantiate our synthetic population with a binary gender (male and female), a generic ethnicity (chosen from 4 levels), a level of physical attractiveness (between 0 and 1), and age (18-45). All agents attributes were assigned uniformly–that is–there is an equal probability of possessing each one of these attributes by any given agent. In both models, we restrict population size to 1000 agents of varying degree of male and female assigned agents. We do not explicitly model agents viewing "profiles" of potential matches but a proxy would be randomization tests embedded in conditional statements. Once executed agents decide to accept/reject the profile uniformly.

The aim of the level 1 (Figure 1) is to emulate the simplest dating app interface. We model this simulation after Tinder, enabling edges within the code to reflect the actions users can take on the app. There are two phases for each agent. Each agent is assigned an average amount of swipes per day and accepts or rejects potential mates that they are shown using a uniform probability distribution, exhausting all their available swipes. When two agents accept each other's profile, an undirected link is created among them, indicating a "match". After all swipes have been exhausted, agents will evaluate all matches acquired, and will select one at random to be messaged. The simulation is then repeated. We set 62% of users as male and 38% as female to emulate proportions found on Tinder (McGrath 2015). Each agent accepts or rejects potential matches based on preferences for age difference, ethnicity-based homophily and attractiveness-based homophily. Males favor females slightly younger than them, and females prefer males slightly older than them (Conway, Noë, Stulp, and Pollet 2015). Based on Byrne et al's (1968) previous work on social attraction theory, we make the assumption that the more one agent prefers another, the more likely they will message the other agent. Multiplicity level 2 (Figure 2) aims to model the interface of an app with similar features to OKCupid. Agents have access to the swipe functions of the lower multiplicity level while being able to send a "like with a message" to increase their match probability.

Generally, every agent possesses a threshold value for 'like score' that is used to evaluate every interaction. Tests (if, then, statements) in Figures 1 and 2 are only activated when the 'like score' of the initiator agent exceeds their base threshold and response actions are also only initiated when said threshold is achieved. This ensures that agents' standards for acceptance of a potential mate are heterogeneous and that agents maintain an individual preference in agent-agent interactions closely representing real life online dating dynamics. Consequently, this process results in a 'like' with increased probability only if the 'like score' exceeds the threshold value and places our model squarely in the mechanistic domain of early threshold models (Granovetter 1978). For example, if the threshold value is not met in the M1 model for males, then the probability of generating a 'like' is 0.3 but if the threshold is met, then the probability is 0.8. All tests are

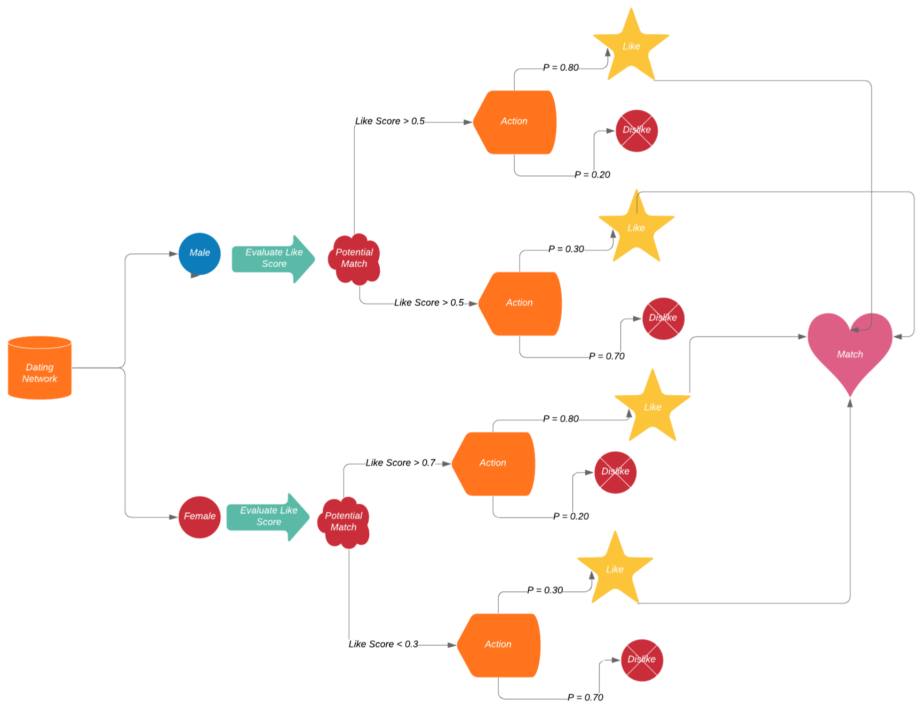


Figure 1: Multiplicity Level One

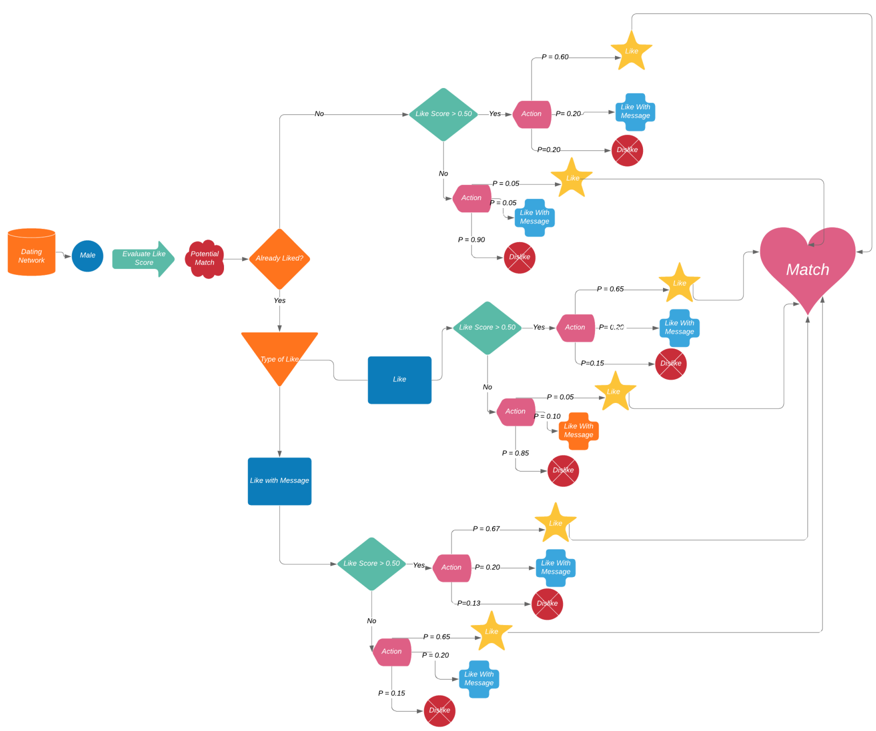


Figure 2: Multiplicity Level Two (Males)

Bernoulli tests (uniform). In the Section 3, we present typical results from a representative run of n=1000 nodes for both model types.

2.4 Internal Validation

Before utilizing our models, we subjected them to rigorous internal validation. Before utilizing our models, we subjected them to rigorous internal validation. Our primary method of validation was to ensure that all synthetic populations were generated to our strict requirements of uniformity for attractiveness [0,1], age [18-40], and ethnicity (categorical) [1,2,3,4]. Both Tables 1 and 1 show the results of a representative run and agrees with our expectations. Since our goal is to quantify the difference of the models based on the given multiplicity of dating apps or the *structure* of the system, no further statistical validation was necessary since the outputs-given-structure is what we hope to determine. However, a number of ad-hoc validation experiment during model development were undertaken to ensure models behaved as intended.

Table 1: Validation of Simulation Inputs for M1 Model

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Attractiveness	1,000	0.521	0.287	0.003	0.282	0.778	0.999
Age	1,000	28.960	6.740	18	23	35	40
Ethnicity	1,000	2.509	1.126	1	2	4	4

Table 2: Validation of Simulation Inputs for M2 Model

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Attractiveness	1,000	0.516	0.294	0.001	0.257	0.769	1.000
Age	1,000	29.057	6.012	19	24	34	39
Ethnicity	1,000	2.446	1.139	1	1	3	4

3 RESULTS

3.1 Gender-Based Distributions of 'Likes', 'Dislikes', 'Matches' and 'Messages'

Figures 1 and 2 summarize the frequency distributions for both M1 (eg. Tinder) and M2 (e.g. OkCupid) model runs. We found similar overall behavior in both M1 and M2 specifically when it comes to the general shape and location of the frequency distributions but there are some telling differences. Both male and female agent in both models liked and dislikes other agents according to some bell-shaped (likely Gaussian) distribution. As we have set the search (profile viewing) mechanism to be random with random (uniform) attributes and set the age evaluation to be a Gaussian difference, this was expected. On average , Female agents liked and disliked more male agents as the simulations were initiated with more male than female agents.[1] While this may be unintuitive, because this model is essentially a matching process, lower counts of female agents mean that they will 'dislike' more but also 'like' more since the male agents posses greater diversity and thus more likely to overcome the female agent's "threshold".

[1] In the Methods and Means section we discussed that we attempted to reflect the state of Tinder in choosing our gender proportions.

However, while female agents (in both models) were more active, the matching distribution did not reflect this activity. The majority of male and female agent received no matches, and while more female agents received 1 matches than their male agent counterparts in both model types, more male agents received 2 or more matches (Figure 3c) or roughly an equal amount (Figure 4c). Both marginal distributions held a right-skew. Finally, the messaging distribution was both skewed and followed similar patterns to the match distribution. Overall, we observed similar stylized facts in our analysis when conditioning on gender between M1 and M2 models.

3.2 Overall Comparison of M1 and M2 Using a Paired t-test

As we hypothesized, the M1 model with its simpler features and focus on generating likes and matches exceeded M2 in scaling activity. Tables 3 and 4 provide summary statistics for both M1 and M2 models. Consider that the average number of likes in M1 was 4 times its M2 counterpart, but that the ratio of matches to likes is 0.05 (5%) and 0.016 (1.6%) for M1 and M2, respectively. Furthermore, because agents in M2 considered more features and more ways to match the number of 'dislikes' reflected this behavior.

Table 3: Summary statistics for M1 Model

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
'Like' Count	1,000	22.449	4.662	10	19	25	40
'Dislike' Count	1,000	55.878	7.360	30	51	61	75
'Match' Count	1,000	1.129	1.152	0	0	2	6
'Message' Count	1,000	0.609	0.917	0	0	1	5

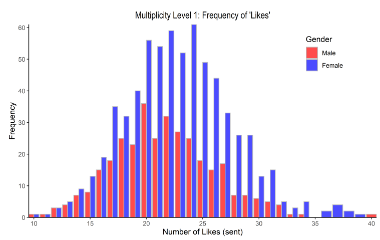
Table 5 summarized the comparison using a paired t-test. We found distinct differences between the models. All results are significant. As shows, M1 (e.g Tinder) exceeded M2 (e.g OkCupid) in producing more likes, matches, and messages, but less dislikes.

Table 4: Summary statistics for M1 Model

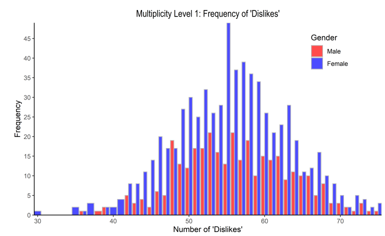
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
'Like' Count	1,000	5.123	2.386	0	3	7	14
'Dislike' Count	1,000	71.141	8.766	45	65	77	100
'Match' Count	1,000	0.302	0.561	0	0	1	3
'Message' Count	1,000	0.083	0.326	0	0	0	3

Table 5: Statistical Comparison ML vs. ML using paired t-test

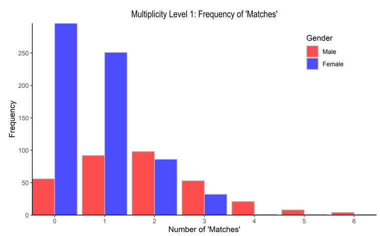
Relation Type	estimate	statistic	p.value	DOF	conf.low	conf.high
'Like'	17.326	107.8	0	999	17.01	17.64
'Dislike'	-15.263	-43.1	2.11e-230	999	-15.957	-14.56
'Match'	0.827	20.4	5.04e-78	999	0.74	0.90
'Message'	0.526	17.0	2.802e-57	999	0.46	0.58



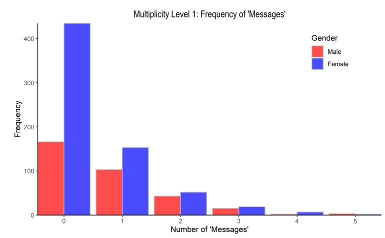
(a) Frequency of 'Likes' for males (red) and females (blue)



(b) Frequency of 'Dislikes' for males (red) and females (blue)

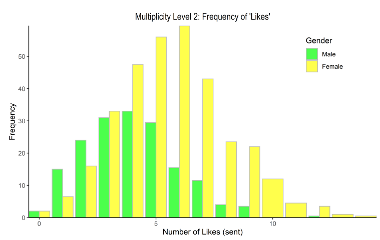


(c) Frequency of 'Matches' for males (red) and females (blue)

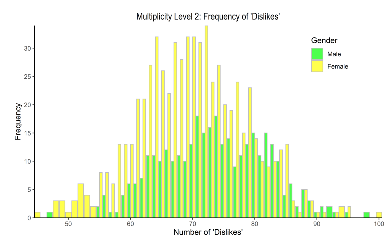


(d) Frequency of 'Messages' for males (red) and females (blue)

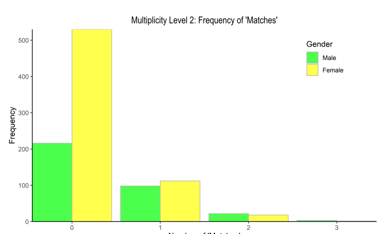
Figure 3: Summary: 'like', 'dislike', 'match', and 'message' distributions for Multiplicity Level 1 models



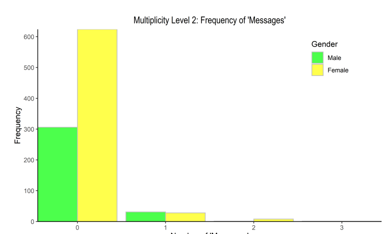
(a) Frequency of 'Likes' for males (green) and females (yellow)



(b) Frequency of 'Dislikes' for males (green) and females (yellow)



(c) Frequency of 'Matches' for males (green) and females (yellow)

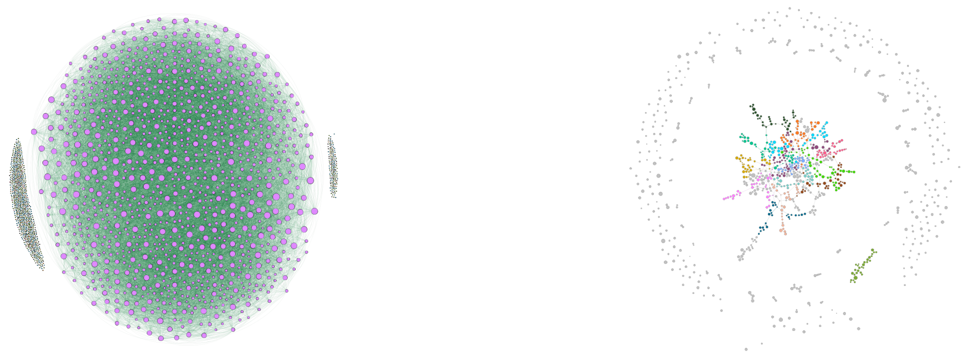


(d) Frequency of 'Messages' for males (green) and females (yellow)

Figure 4: Summary: 'like', 'dislike', 'match', and 'message' distributions for Multiplicity Level 2 models

3.3 Network Results and Regression Model

We also issued a direct comparison between the M1 and M2 models resultant 'like' and 'match' networks shown in Figures 5 and 6. As might be expected, since much of the mechanism of the underlying rules was set to follow random behavior, the like networks of both contained little non-random structure (by visual inspection). Since M1 generated more activity than M2, the density of M1 was higher than M2. More interestingly, the match networks for both models did contain non-random structure (figure 5b and 6b). In the M! model the match network produced a large component and many isolates, suggesting that simple rules can emerge match networks with structures indicative that small preferences (in age for example) can produce clusters of high matching probability. The M2 match network showed highly dyadic structure (rather than transitive) suggesting that agents matched with one or two other agents but not within larger clusters of agents as in M1. We remind the reader that that the main difference between behaviors in our models is simply that in M2 the user considers more features.



(a) M1 'Like' Network

(b) M1 'Match' Network

Figure 5: M1 cross-sectional resultant networks for a single run including isolates

(a) M2 'Like' Network

(b) M2 'Match' Network

Figure 6: M2 cross-sectional resultant networks for a single run including isolates

Aside from summary network statistics and visual model inspection, we utilized the network regression technique, Multiple Regression Quadratic Assignment Procedure (MRQAP) (Mantel 1967) to investigate the relationship of our monadic and dyadic covariates to the production of a match. MRQAP is best used on dyadic relations and so in order to utilize it effectively we converted out monadic covariates, such as attractiveness, age, and ethnicity into difference adjacency matrices. We also included the 'like' and 'dislike' networks as independent variables. Our question, simply put, is what are the effects of our covariates on producing a match, given the multiplicity of our two models. Table 6 provides statistically significant models considered over 100 iterations that answer this question.

Table 6: MRQAP Linear Regression Comparison Using Monadic and Dyadic Covariates

	Dependent Network: 'Match' Network	
	Multiplicity Level 1	Multiplicity Level 2
'like' relation	$-1.227880e-03$	$-7.184141e-02$
Pr(<=b)	(0.0)	(0.0)
Pr(>=b)	(1.0)	(1.0)
Pr(>=|b|)	(0.0)	(0.0)
'dislike' relation	$-1.227854e-03$	$-7.183297e-02$
Pr(<=b)	(0.0)	(0.0)
Pr(>=b)	(1.0)	(1.0)
Pr(>=|b|)	(0.0)	(0.0)
attractiveness (difference)	$-6.221186e-05$	$-6.871735e-04$
Pr(<=b)	(0.07)	(0.37)
Pr(>=b)	(0.93)	(0.63)
Pr(>=|b|)	(0.19)	(0.8)
age (difference)	$-5.836387e-06$	$-5.691259e-05$
Pr(<=b)	(0.0)	(0.32)
Pr(>=b)	(1.0)	(0.68)
Pr(>=|b|)	(0.0)	(0.59)
ethnicity (difference)	$1.629981e-04$	$-5.710848e-04$
Pr(<=b)	(0.98)	(0.22)
Pr(>=b)	(0.02)	(0.78)
Pr(>=|b|)	(0.08)	(0.4)
intercept	$1.104061e-03$	$7.226399e-02$
Pr(<=b)	(1.0)	(1.0)
Pr(>=b)	(0.0)	(0.0)
Pr(>=|b|)	(0.0)	(0.0)
Replication(s)	1,00	1,00
R^2	0.0001038	0.0006752
Adjusted R^2	9.88e-05	0.0006702
Residual Std. Error (df = 998994)	0.0336	0.2571
F Statistic (df = 5; 998994)	20.74	135
Note:	*both models are significant at the* p<0.0001 *level*	

As might be expected almost all parameter estimates are both small and negative. This is due to obvious factors, primarily, since to ensure a proper baseline of comparison, we set the majority of our randomization test to follow uniformity–this, by-design–would ensure that parameter estimates are small. If we were to use real world input distributions for frequency of likes, dislikes and others, we may find greater effect sizes.[2] However, the difference in order of magnitude for each parameter estimate is telling on which effects

[2] However, this data is not made available publicly and is considered highly proprietary.

contribute more to the production of match edges. Secondly, while parameter estimates may be small their negative sign (with exception to ethnicity) signifies the reduced density of the match network when compared to the like network. Thus, both models presented in table 6 seem to possess parameter estimates that pass out litmus test. Finally, since MRQAP does not produce traditional p-values we have included the proportion of simulated networks that meet lower-end one-tailed, upper-end one-tailed, and two-tailed tests. We will proceed in explaining and interpreting each effect.[3]

In both M1 and M2 models, liking and disliking produces statistically significant effects. In fact, for 100 iterations considered by MRQAP, no networks produced an effect as large as the ones presented (one-tailed test) in the table for the 'like and 'dislike' effects. Interestingly, the magnitude of the effect was an order larger for M1 than it was for M2, suggesting that 'liking' and 'disliking' has a greater influence matching dynamics in M1 than in M2. Attractiveness difference and age difference followed a similar pattern, with the M1 effect influencing matching more than the M2 effect. However, these effects were 2 and 3 orders of magnitude smaller than 'like' and 'dislike' effects. That is, in both models 'like' relations are much more closely related to who you ultimately match with than age difference or attractiveness difference. Finally, ethnicity (difference) for both models was a positive effect and 98% of iterations were below this effect - in essence this covariate is significant on the lower-tail test, and since its sign is positive, then it reveals that *similarity*, not difference is what plays a role; otherwise known as homophily. That is, the more similarity exists between agents engaged in a like interaction the more likely a match will occur. Also of interest is the size of this effect which is roughly 3.5 times larger for M2 than for M1.

4 DISCUSSION

Both models were statistically significant, though that evidence in isolation is simply indicative of the dependencies we built into the model in terms of agent-based rules. However, effect sizes considered in light of the number of options available to agents to achieve a match (multiplicity) are instructive. Consider that for the M1 model the order of importance by effect size is 'like' relation, 'dislike' relation (an analogue to 'like'), ethnic similarity, attractiveness difference, and finally age difference. Given the same effects in the network regression model, the ordering for M2 diverges from this with 'like' and 'dislike', and age difference occupying the same positions with an order (or more) of magnitude difference. In M2, ethnicity similarity and attractiveness difference are of roughly the same magnitude and certainly the same numerical order. This is direct evidence supporting our hypothesis–that holding all else equal, multiplicity allows for more pathways in achieving a 'match' and consequently, agents must consider those new pathways. In so doing social effects are transformed. In our specific case, the importance of attractiveness increased to the level of ethnicity, and earlier we showed that even the total number of matches differed in a significant way from both simulations.

5 CONCLUSIONS

In this paper, we have shown, through agent-based simulations how additional features (multiplicity) in online dating applications can cause differences in overall matches and how the underlying social effects can be transformed by the those same features. There are many directions we could take this area of inquiry, including calibrating our social effect, models, adding weighted dyadic relationships and most importantly using some reference data set The later represents our work's greatest limitation. We hope to be able to pursue these avenues in future papers.

[3]As a reminder, our dependent variable is the match network, and our independent variables are the like network, dislike network, monadic attractiveness converted to dyadic difference in attractiveness, monadic age converted to dyadic age difference, and monadic ethnicity converted to dyadic difference in ethnicity.

REFERENCES

Conway, J. R., N. Noë, G. Stulp, and T. V. Pollet. 2015. "Finding your Soulmate: Homosexual and heterosexual age preferences in online dating". *Personal Relationships* vol. 22 (4), pp. 666,678.

Fiore, A., and J. Donath. 2005. "Homophily in online dating: when do you like someone like yourself?". In *CHI '05 Extended Abstracts on human factors in computing systems*, CHI EA '05, pp. 1371,1374, ACM.

Granovetter, M. 1978. "Threshold Models of Collective Behavior". *The American Journal of Sociology* vol. 83 (6), pp. 1420–1443.

Greenlees, I., and W. Mcgrew. 1994. "Sex and age differences in preferences and tactics of mate attraction: Analysis of published advertisements". *Ethology and Sociobiology* vol. 15 (2), pp. 59,72.

Koch, Lucy 2019. "Dating App Growth Slows, but Advertisers Shouldn't Ignore Hopeful Singles". Retrieved from https://www.emarketer.com/content/dating-app-growth-slows-but-advertisers-shouldnt-ignore-hopeful-single-users.

Krackardt, D. 1987. "QAP partialling as a test of spuriousness". *Social Networks* vol. 9 (2), pp. 171,186.

Mantel, N. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach". *Cancer Research* vol. 27 (2 Part 1), pp. 209–220.

McGrath, Felim 2015. "What to Know About Tinder in 5 Charts". Retrieved from https://blog.globalwebindex.com/trends/what-to-know-about-tinder/, accessed 2019-09-25.

McPherson, M., L. Smith-Lovin, and J. M. Cook. 2001. "BIRDS OF A FEATHER: Homophily in Social Networks". *Annual Review of Sociology*.

Sumter, S. R., L. Vandenbosch, and L. Ligtenberg. 2017. "Love me Tinder: Untangling emerging adults' motivations for using the dating application Tinder". *Telematics and Informatics* vol. 34 (1), pp. 67–78.

Tyson, G., V. Perta, H. Haddadi, and M. Seto. 2016. "A First Look at User Activity on Tinder". *arXiv.org*.

AUTHOR BIOGRAPHIES

LIAM MCKENNA is a senior at George Mason University is approaching his final semester majoring in Computational and Data Science and minoring in Data Analysis. He worked as a data science intern this past summer at Noblis and plans to return to work full time after graduation. His email is lmckenna@gmu.edu.

FATMA ABDULWAHAB is a graduating senior completing a Bachelor of Science in Computational and Data Science with a double minor in Mathematics and Sociology at George Mason University. Her email address is fabdulwa@gmu.edu.

STEVEN HOANG is a junior at George Mason University pursuing a Bachelor of Science in Computational and Data Science with a minor in Biology. He is currently working as a Data Analyst for the Vietnamese Boat People Podcast. His email address is shoang@gmu.edu

COLLIN HENLEY is an upcoming college senior, is working on his Bachelor's in Computational and Data Science with a minor in statistics at George Mason University. Most recently, he worked as an Operations Research Analyst at Fort Belvoir. His email address is chenley4@gmu.edu.

JOSEPH A.E. SHAHEEN, PHD is the ORISE Intelligence Community Postdoc Fellow and Adjunct Professor at George Mason University. His email address is jshaheen@gmu.edu.