

# Der Einfluss von Matching-Algorithmen auf den Erfolg einer Dating-App

Lucas Friedrich Mertens<sup>1</sup>, Erik Christian Müller<sup>2</sup>, Johanna Sebastian<sup>3</sup>

Freie Universität, Berlin, Deutschland

<sup>1</sup>lmertens@zedat.fu-berlin.de, <sup>2</sup>erikcm@zedat.fu-berlin.de,

<sup>3</sup>jsebastian@zedat.fu-berlin.de

**Abstract.** Für den Erfolg von Online-Dating-Plattformen sind Matching-Algorithmen, die für passende, zielorientierte Partnervorschläge verantwortlich sind, von übergeordneter Bedeutung. Fokus dieser Arbeit ist die Betrachtung einer differenzierten Maximierung der wahrgenommenen Zufriedenheit der gesamten Population einer Dating App in Abhängigkeit von ihren Motiven zur Anmeldung auf dieser Plattform. Es wurden verschiedene Matching-Algorithmen implementiert und deren Erfolg verglichen. Es konnte festgestellt werden, dass ein *Preferred Proposing* (individualisierte Partnervorschläge) gegenüber zufälligen Partnervorschlägen zu einer maßgeblichen Steigerung der wahrgenommenen Zufriedenheit der Population führt. Weiterhin konnte die negative Auswirkung von Werbung und Fakeprofilen innerhalb der Plattform quantifiziert werden. Die stochastische agentenbasierte Simulation wurde mit *AnyLogic PLE* durchgeführt. Auf dieser Grundlage wurde das Matching-Verhalten der heterogenen Population visualisiert. Mit Hilfe dieser Visualisierung lässt sich das Verhalten der gesamten Population global sowie auf Ebene des Individuums präzise betrachten und nachvollziehen.

**Keywords:** Agent-based Simulation, Online Dating, Stochastic Simulation

## 1 Einleitung

Online-Dating-Plattformen sind in der Bevölkerung beliebt und werden in der Wissenschaft zunehmend untersucht [1], [2]. Rund zehn Prozent der Deutschen sind aktive Nutzer von Online-Dating-Plattformen [3] und circa ein Drittel der Bevölkerung hatten bereits ein Date über eine solche Plattform [4]. Bekannte Beispiele sind *Parship*, *Tinder* und *Elite-Partner*. Um sich an diesem umkämpften Markt langfristig durchzusetzen, ist es wichtig, eine möglichst zufriedene Nutzerbasis aufzubauen und aufrechtzuerhalten. Dafür ist vor allem ein erfolgreicher Matching-Algorithmus nötig. Matching-Algorithmen basieren auf den durch Vorschlagssysteme automatisch generierten Empfehlungen. Sie helfen den Nutzern die Vielzahl an Möglichkeiten und die überwältigende Menge an Informationen drastisch zu reduzieren [5].

Bestehende Arbeiten zu Methoden, auf denen Vorschlagssysteme basieren, lassen sich dabei in unterschiedliche Strömungen unterteilen. Eine Strömung nutzt Gemeinsamkeiten der Nutzer, um diese zu clustern [5], [6]. Dabei werden explizite und

implizite Daten genutzt [6]. Eine zweite Strömung verwendet Data-Mining-Techniken [8] und eine weitere Machine-Learning-Techniken [9]. Während diese Arbeiten das Wissen über unterschiedliche Methoden für Matching-Algorithmen erweitert haben, fehlt derzeit eine detaillierte Untersuchung der Effekte von Matching-Algorithmen auf die Zufriedenheit der Nutzer. Außerdem ist ein besseres Verständnis der Funktionsweise von Matching-Algorithmen anzustreben, um herauszufinden, wie diese für ein effizientes Matching und eine möglichst glückliche Population justiert werden können.

*Das Ziel dieser Arbeit besteht darin, das Verhalten einer heterogenen Population zu analysieren, visualisieren und einen Matching-Algorithmus zu entwerfen, der die Matching-Quote und das Glück der Population maximiert.* Die Matching-Quote spiegelt dabei die relative Anzahl von Matches im Verhältnis zu den gelikten Nutzern wider. So ist davon auszugehen, dass ein hohes Verhältnis von Likes zu Matches in einem positiven Nutzungserlebnis resultiert und einen erfolgreichen Matching-Algorithmus darstellt. Dabei ist insbesondere eine vielfältige Basis von Nutzern abzubilden und zu berücksichtigen. Dies betrifft sowohl deren individuelle optische Ausprägungen, als auch deren Wünsche und pluralistischen Motive für die Nutzung einer Dating-Plattform [10]. Gewisse Ereignisse wie Matches oder der Erhalt von Nachrichten werden von jedem Nutzer in Abhängigkeit zu diesen Motiven differenziert wahrgenommen. So nimmt ein Nutzer z.B. den Erhalt eines Matches anders wahr, wenn er eine Beziehung anstrebt, als jemand der nach Selbstbestätigung sucht. Um diese Vielfalt abzubilden, nutzt diese Arbeit eine agentenbasierte Simulation [10], [11]. Diese Methode ist geeignet, da sie sich insbesondere auf die einzelnen aktiven Komponenten (Agenten) eines Systems konzentriert und diese unabhängig abbilden kann. In agentenbasierten Simulationen wird das Verhalten der Agenten (hier die Nutzer) klar strukturiert definiert. Verbindungen zwischen ihnen werden aufgebaut, Umgebungsvariablen gesetzt und Simulationen ausgeführt. Die globale Dynamik des Systems entsteht aus den Interaktionen einer Vielzahl einzelner Verhaltensweisen.

## 2 Literaturrecherche

Matching-Algorithmen basieren auf den durch Vorschlagssysteme automatisiert generierten Empfehlungen. Mit dem Aufkommen von Social-Matching-Systemen [13] stehen vor allem reziproke Vorschlagssysteme im Fokus der Forschung [14]. Beispiele hierfür sind Jobbörsen und das Online-Dating. Diese sog. Zwei-Wege-Systeme [5], [6], [14] werden genutzt, um zwei Personen einander vorzuschlagen [16]. Dabei entsteht nur eine erfolgreiche Verbindung, wenn sich die Wünsche und Präferenzen beider Personen überschneiden [17]. Aufgrund der großen Zahl aktiver Nutzer einer Online-Dating-Plattform, ist es für Nutzer unmöglich den richtigen Partner ohne die Hilfe von Algorithmen herauszufiltern [4], [14]. Hierbei können reziproke Vorschlagssysteme den Nutzer unterstützen [18]. Diese beruhen auf expliziten sowie impliziten, durch die Verhaltensweisen ermittelten, Informationen [3], [4], [12], [17]. Explizite Informationen bestehen meist aus demografischen Inhalten, einem Foto, den Präferenzen bzgl. des Partners und einem kurzen Text über die Person selbst [16], [17].

Um diese Vielzahl an Informationen für ein Vorschlagssystem zu verarbeiten, gibt es in der Literatur unterschiedliche Untersuchungsstränge.

Ein Hauptstrang ist das Filtern nach Gemeinsamkeiten der Nutzer anhand des Profils und Verhaltens auf der Plattform [5], [6]. Je mehr Faktoren für das Auffinden von Gemeinsamkeiten in Betracht gezogen werden, desto erfolversprechender sind die Vorschläge [19]. So nutzen beispielsweise Chen et al. [15] explizite sowie implizite Informationen, um die Gemeinsamkeiten zweier Nutzer in einem Cluster genauer zu berechnen. Akehurst et al. [20] bestätigen, dass Nutzer mit ähnlichen Profilen vergleichbare Personen liken. Zusätzlich lassen sich Informationen über vergangene Beziehungen der Nutzer mit einbeziehen [21].

Eine weitere Möglichkeit besteht darin, die Gemeinsamkeiten der Nutzer anhand der Ähnlichkeit ihrer Präferenzen und vergleichbarer Attraktivität zu identifizieren [7]. Eine ähnliche Präferenz ergibt sich aus dem Versenden von Nachrichten an dieselbe Nutzergruppe und eine vergleichbare Attraktivität aus dem Erhalt von Nachrichten einer homogenen Nutzergruppe [7]. So entwickeln Li und Li [22] ein generelles Framework basierend auf den Gemeinsamkeiten der Nutzung unter Betrachtung des gesamten Online-Dating-Netzwerks. Krzywicki et al. [5] zeigen einen zweistufigen Vorschlagsalgorithmus, in dem zunächst nach Gemeinsamkeiten gefiltert und später die Vorschläge anhand eines Entscheidungsbaumes neu gerankt werden.

Ein zweiter Untersuchungsstrang basiert auf Data-Mining-Techniken, um die Qualität der Vorschläge zu erhöhen [8]. Eine dieser Techniken ist das Clustering, das sowohl die Anzahl, als auch die Komplexität der Daten in einem Matching-Prozess reduziert [23]. Clustering-Techniken unterstützen zudem den Matching-Prozess, indem bestehende Probleme wie das Kaltstart-Problem oder die Seltenheit der Daten gelöst werden [13]. Nutzer lassen sich beispielsweise in männliche und weibliche Cluster einordnen, um diese einander vorzuschlagen. Bei nicht ausreichend personalisierten Vorschlägen, werden jedoch jedem Nutzer eines Clusters dieselben Kandidaten empfohlen [13].

Ein dritter Ansatz basiert auf Machine-Learning-Techniken [24]. Xia et al. [9] nutzen diese, um Vorhersagen darüber treffen zu können, ob ein Nutzer auf eine Nachricht antworten wird oder nicht. Dies unterstützt Vorschlagssysteme dabei, negative Erfahrungen bei initialen Kontaktaufnahmen zu vermeiden, da diese meist mit hohen Kosten für die Plattform verbunden sind. In der Literatur wird eine erfolgreiche Verbindung als Interaktionen (Austausch von Nachrichten) zwischen zwei Nutzern definiert. Eine Möglichkeit den Matching-Algorithmus zu evaluieren, sind dessen Auswirkungen auf die Matching-Quote [18]. Zusätzlich ist aufgrund der großen Anzahl an Nutzern und Informationen die Skalierbarkeit eine wichtige Eigenschaft eines Matching-Algorithmus.

Es lässt sich erkennen, dass in der bisherigen Literatur unterschiedliche Methoden entwickelt worden sind, um Vorschläge von höherer Qualität für die Nutzer einer Online-Dating-Plattform zu generieren. Dazu werden meist historische oder Realdaten verwendet um mit neuen Methoden eine erhöhte Erfolgsrate zu erreichen [4], [14], [16]. Als Baseline wird dabei meist die Erfolgsrate der vorherigen Methode genutzt [5], [14], [17]. Alternativ werden die Ergebnisse verschiedener Methoden verglichen [3], [13]. Ein Vergleich zu einer Baseline mit zufälligen Vorschlägen liegt bisher nicht vor.

Zudem steht in bisherigen Arbeiten der Erfolg der Methode im Vordergrund. Darüber hinaus ist in dieser Arbeit die Erhöhung der Zufriedenheit der heterogenen Nutzerbasis anzustreben [1]. Diese ist vom Motiv zur Anmeldung auf der Plattform abhängig. Welche Motive die Nutzer verfolgen und wie sich diese auf das Matching-Verhalten auswirken, wurde bislang nicht intensiv untersucht. Weiterhin zeigt die Untersuchung der Literatur, dass einige Methoden zwar wissenschaftlich ausgearbeitet und dokumentiert wurden, es jedoch an Visualisierungen der Methoden für die Praxis fehlt. Dies ist der Ansatzpunkt der vorliegenden Arbeit.

### 3 Methodisches Framework

#### 3.1 Zielstellung und Aufbau des Modells

Die Beschreibung des konzeptionellen Modells der Simulation orientiert sich am Sieben-Schritte-Ansatz von Law [25] sowie dem ODD-Protokoll von Grimm et al. [26].

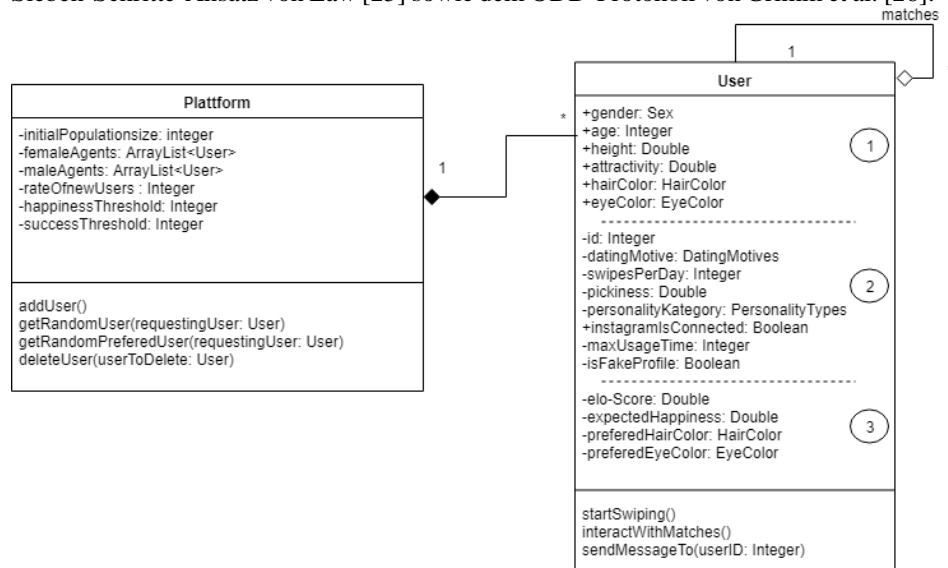


Abbildung 1. Statisches Systemverhalten

Ziel des Modells ist es, das Verhalten von Nutzern einer Online-Dating-Plattform abzubilden und die Auswirkung verschiedener Matching-Algorithmen darzustellen. Das Modell in der vorliegenden Arbeit orientiert sich an der Dating-Plattform *Tinder*. Für eine möglichst realitätsnahe Betrachtung der Interaktionen von Plattform zu Nutzer und Nutzer zu Nutzer werden in *Abbildung 1* die in der Simulation betrachteten relevanten Attribute und Methoden der einzelnen Akteure dargestellt. Die Nutzer der Plattform werden durch persönliche, plattformunabhängige Attribute (1) und plattformbasierte Attribute charakterisiert (2,3). Initialisiert werden die plattformunabhängigen Attribute anhand der Profilbild-Studie 2019/2020 [27]. Die plattformbasierten Attribute (2,3) werden gleichverteilt initialisiert und unterteilen sich

weiterhin in statische (2) und dynamische (3) Attribute. Statische Attribute bleiben während der Simulation fixiert, wohingegen sich dynamische Attribute, wie z.B. die vermutete Zufriedenheit, im Verlauf verändern.

Für die Messung der Zufriedenheit von Nutzern ist das Motiv ihrer Anmeldung eine maßgebliche Voraussetzung. Für die implementierte Simulation werden basierend auf Menkin et al. [10] die folgenden vier Motive angenommen: Amusement, Ego-Boost, Affair, Relationship. Für jedes dieser Motive führt ein anderes Ereignis zu einer maßgeblichen Veränderung der Zufriedenheit:

1. Amusement: es wird eine möglichst große Nutzerbasis benötigt, um kontinuierlich Vorschläge zu ermöglichen und ein dauerhaftes Entertainment zu gewährleisten.
2. Ego-Boost: der Nutzer strebt eine hohe Anzahl von Matches an, um sein Selbstwertgefühl zu erhöhen.
3. Affair: es werden möglichst viele Matches mit anschließendem Nachrichtenfluss gewünscht, um dem Nutzer die Möglichkeit zu geben, eine Affäre zu beginnen.
4. Relationship: es werden möglichst viele Matches mit Nutzern gewünscht, die das gleiche Motiv verfolgen und ein anschließender Nachrichtenfluss angestrebt.

Neben den Agenten (Nutzern) bettet das konzeptionelle Modell die Plattform mit Eigenschaften wie der Anzahl an Nutzern zum Start der Simulation oder einer Zuwachsrate an Nutzern während eines Simulationsschrittes ein. Simuliert wird ein Monat (31 Tage) mit je drei Onlinezyklen (Steps). Gemäß dem initialisierten Online-Verhalten kann ein Nutzer in einem, zwei oder in allen drei Zyklen online sein.

### 3.2 Dynamisches Systemverhalten

In jedem Onlinezyklus eines Agenten durchläuft dieser verschiedene Phasen, welche vergleichbar mit dem Verhalten von Nutzern der Dating-App *Tinder* sind. Im ersten Schritt jedes Zyklus überprüft der Agent seine vorhandenen Matches. Sollte er noch keine Matches haben, besucht er das Swipe-Menü und eruiert mehrere Nutzer anhand ihrer wahrgenommenen Attraktivität und entscheidet sich für oder gegen ein Like. Sollte er daraufhin Matches erhalten oder bereits zuvor Matches erhalten haben, entscheidet er sich für oder gegen das Verschicken einer Nachricht an sein Match. Sollten zwei Nutzer in gegenseitigen Kontakt treten und eine Beziehung suchen, werden sie sich im Laufe der Zeit von der Plattform abmelden. Sollte der Kontakt (eine Nachricht) nicht erwidert werden, wird der Nutzer stattdessen das Match auflösen. Im Anschluss an diese Interaktionen wird der Nutzer weiter das Swipe-Menü besuchen.

Je nach ausgelöstem Ereignis passt sich im Laufe der Simulation die erwartete Zufriedenheit (*Expected Happiness*) und der Elo-Score der Nutzer an. Der Elo-Score eines Nutzers spiegelt hierbei seine Like-Wahrscheinlichkeit wider. Je höher der Elo-Score eines Nutzers ist, desto eher wird dieser gelikt. Der Elo-Score wird in diesem Modell ausschließlich durch das Like-Verhalten der Population angepasst und dadurch beeinflusst, ob und von wem ein Nutzer ein Like erhält. Aktivitäten, welche die *Expected Happiness* erhöhen, sind das Entstehen eines Matches und der Erhalt einer Nachricht. Im Gegensatz dazu wird erwartet, dass ein Nutzer unglücklicher wird, wenn er einen Nutzer likt und kein Match entsteht, er ein Fakeprofil matcht, Werbung sieht oder es keine neuen Nutzer mehr in der Plattform gibt. Die Quantifizierung des

Einflusses dieser positiven und negativen Ereignisse hängt dabei von den Motiven des Nutzers ab. So wird beispielsweise davon ausgegangen, dass ein Nutzer, der auf der Suche nach Selbstbestätigung ist, den Erhalt eines Matches weitaus höher gewichtet, als ein Nutzer, der die App zum Zeitvertreib nutzt. Um das konzeptionelle Modell zu validieren, ist ein Walkthrough anhand der bestehenden Applikation der Online-Dating-Plattform *Tinder* vorgenommen worden [25]. Weiterhin wurde das Modell mit zwei wissenschaftlichen Betreuern innerhalb von mehreren Sessions durchgegangen und mit einer Mitarbeiterin einer Online-Dating-Plattform besprochen. Dabei wurde jeder Schritt der Simulation erneut einzeln analysiert und auf Realisierbarkeit sowie Korrektheit überprüft. So konnten vor Beginn der Implementierung konzeptionelle Probleme schnell und effizient entdeckt und verbessert werden.

## 4 Implementierung

Das konzeptionelle Modell wird mit Hilfe einer agentenbasierten Simulation [25] implementiert. Zur Umsetzung dieses Konzepts wurde *AnyLogic PLE* gewählt [28]. Kernbestandteil der Implementierung ist es, das individuelle Nutzerverhalten während eines Simulationsschrittes und dessen atomaren Bestandteilen möglichst realitätsnah abzubilden. Jeder Simulationsschritt eines Nutzers unterteilt sich dafür gemäß dem im 3.2 *Dynamisches Systemverhalten* dargestellten Ablauf in zwei Hauptbestandteile: das Matchen von Nutzern und deren Interaktion.

### 4.1 Der Matching-Algorithmus

Für ein erfolgreiches Matchen stehen gezielte Partnervorschläge im Mittelpunkt. Dieser Prozess unterteilt sich auf Nutzer-Ebene in den Erhalt von Vorschlägen und dem Auswertungsprozess, ob dieser gelikt wird. Für das Vorschlagen neuer potenzieller Matches werden im implementierten Modell als Baseline zufällig Nutzer vorzuschlagen. Wie in 6.2 *Numerische Ergebnisse* genauer dargestellt, führt dies zu einer unbefriedigenden Benutzererfahrung. Um die Matching-Wahrscheinlichkeit zu erhöhen, wurde anhand des beobachteten Verhaltens der Agenten alternativ ein *Preferred Proposing* (gezieltes Vorschlagen) anderer Nutzer implementiert. Dafür gilt es die Like-Quote (Likes/Vorschläge) und insbesondere die Matching-Quote (Matches/Likes) zu maximieren. Es ist nicht zielführend, jedem Nutzer das attraktivste Profil der Plattform vorzuschlagen. Sie werden dieses zwar liken, jedoch wird bei oberflächlichen Dating-Apps nur ein sehr geringer Prozentsatz des anderen Geschlechts das Match erhalten. Im Vergleich zum zufälligen Vorschlagen werden deshalb beim *Preferred Proposing* erst alle Nutzer vorgeschlagen, die den Nutzer gelikt haben und deren Elo-Score nicht zu weit abweicht. Somit wird erreicht, dass ähnlich attraktive Nutzer eine höhere Wahrscheinlichkeit haben ein Like zurück zu bekommen. Sollte der aktive Nutzer keine Likes von anderen ähnlich attraktiven Nutzern mehr haben, ist die Like-Wahrscheinlichkeit im weiteren Auswahlprozess zu maximieren. Dabei wird stets darauf geachtet, dass eine realistische Matching-Wahrscheinlichkeit der Vorschläge vorliegt. Im ersten Schritt werden beim *Preferred Proposing* Nutzer vorgeschlagen, die einen ähnlichen Elo-Score haben und der bevorzugten Augen- bzw. Haarfarbe des

aktiven Nutzers entsprechen. Die Wahl von Augen- und Haarfarbe als subjektives Präferenzkriterium erfolgte bei der Implementierung willkürlich und dient als Mittel zur Umsetzung der Präferenzfassung und -messung. Sollte es keine weiteren Agenten in der Population geben, welche den subjektiven Präferenzkriterien entsprechen, folgt eine iterative Auflockerung der Kriterien. So wird in der ersten Iteration die Bedingung der bevorzugten Präferenzen aufgelockert und anschließend die Bedingung eines ähnlichen Elo-Scores aufgehoben. Sollte es durch ein exzessives Verwenden der Dating-App keine weiteren Nutzer mehr geben, welche diesen Eigenschaften entsprechen, werden zufällige Nutzer vorgeschlagen.

Drei Faktoren beeinflussen die Like-Wahrscheinlichkeit eines vorgeschlagenen Profils: wie wählerisch der Nutzer ist (*Pickiness*), wie stark seine Präferenzen bzgl. der Haar- und Augenfarbe des vorgeschlagenen Nutzers sind sowie einem Bonusfaktor. Die *Pickiness* unterliegt einer Normalverteilung mit einem Mittelwert von 0.5 in dem Intervall von null bis eins. Autark betrachtet würde eine *Pickiness* von 0.5 in einem Like der Hälfte aller vorgeschlagenen Nutzer resultieren und bei eins würde jeder Vorschlag gelikt werden. Die gewichtete Augen- und Haarfarbenpräferenz hängt von dem vorherigen Swipe-Verhalten des Nutzers ab und schwankt zwischen null und eins. Das Resultat dieser Faktoren wird addiert und durch die Anzahl der Summanden (hier drei) dividiert. Der maximale Wert dieses Terms könnte eins annehmen und wäre als 100%iger Like zu interpretieren. Weiterhin wird die Like-Wahrscheinlichkeit über einen Bonusfaktor beeinflusst. Dieser Faktor erhöht die Like-Wahrscheinlichkeit, wenn der vorgeschlagene Nutzer einem ähnlichen Elo-Score unterliegt, einen höheren Attraktivitätswert besitzt oder seinen Social-Media-Account mit seinem Profil verknüpft hat.

#### **4.2 Interaktion mit Matches**

Hat ein Nutzer bereits Matches, überprüft er in jedem Simulationsschritt ob er mit diesen in Kontakt treten möchte. Abhängig davon welchem Geschlecht der Nutzer entspricht, welches Datingmotiv er verfolgt und ob er bereits eine Nachricht von dem gematchten Nutzer erhalten hat, entscheidet er sich zu einer spezifizierten Wahrscheinlichkeit für oder gegen das Schreiben einer Nachricht an sein Match. Der Einfluss der Parameter auf diese Wahrscheinlichkeiten basiert dabei auf getroffenen Annahmen und Erfahrungswerten. So wird beispielsweise davon ausgegangen, dass es viel wahrscheinlicher ist, dass der Mann den ersten Schritt wagt und seinem weiblichen Match die erste Nachricht schreibt. Bekommt der gematchte Nutzer eine Nachricht, erhöht sich, wie später genauer dargestellt, seine *Expected Happiness*. Sollten sich beide gematchten Nutzer eine Nachricht geschrieben haben, resultiert das in einem Chat und in der Implementierung wird implizit von einer Zielerreichung der Agenten ausgegangen. Verfolgen beispielsweise beide Agenten das Ziel *Relationship* oder *Affair* wird von einem erfolgreichen Finden einer Beziehung bzw. Affäre ausgegangen. Im ersten Fall melden sich beide Nutzer sofort von der App ab, wohingegen Nutzer, welche eine Affäre suchen, sich erst nach mehrmaligem Erfolg von der App abmelden.

## 5 Visualisierung

Neben der Implementierung verschiedener Matching-Algorithmen, steht auch die Visualisierung der Simulation im Mittelpunkt dieser Arbeit. Für eine geeignete Darstellung sollte es möglich sein, einzelne Individuen identifizieren zu können und möglichst schnell, viele Aussagen über ihr Dating-Verhalten und Nutzungserlebnis zu treffen.

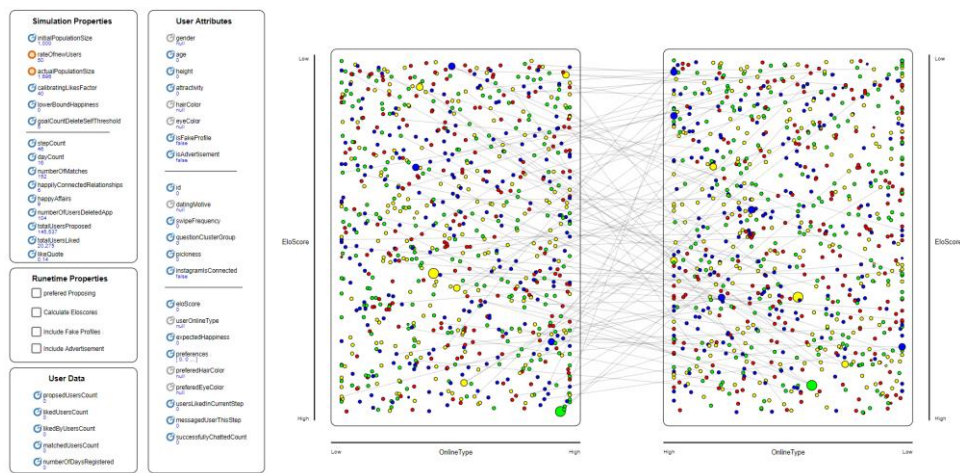


Abbildung 3. Visualisierung der Dating-App

Auf der rechten Seite in *Abbildung 3* ist die gesamte Population zu sehen. Dabei sind männliche Nutzer dem linken und weibliche dem rechten Pool zugeordnet. Bei einer initialen Population existieren somit circa 500 Nutzer je Pool. Farblich wird dabei eine Unterteilung nach den Motiven vorgenommen. Gelbe Individuen sind auf der Suche nach einem Ego-Boost, Grüne auf der Suche nach einer Beziehung, Rote suchen eine Affäre und Blaue sind aus Entertainmentgründen/Langeweile angemeldet. Die Verteilung auf der X-Achse erfolgt nach dem initialisierten Onlineverhalten. Je öfter ein Nutzer online ist, desto dichter sind sie dem Pool des anderen Geschlechts angeordnet. Die Verteilung auf der Y-Achse entspricht dem Elo-Score des Nutzers. Zu sehen ist die Simulation nach einer Woche. Ersten Verbindungen zwischen Nutzern sind entstanden und werden durch Verbindungen/Linien gekennzeichnet. Da ein Nutzer mehrere Matches haben kann, können auch mehrere Linien von ihm abgehen. Je größer der Kreis des Individuums ist, desto mehr Verbindungen gehen von ihm aus. Um die Aussagekraft dieser Übersicht zu erhöhen, wurde sie auf der linken Seite um ein Datendashboard erweitert. Das Dashboard gibt Informationen zu der Simulation allgemein und bei Bedarf über einen beliebigen Agenten während der Simulation wider. Unterteilt wird im Dashboard nach Eigenschaften der Simulation und eines gewählten Nutzers. Dargestellte Eigenschaften der Simulation sind beispielsweise die initiale Populationsgröße oder die Zuwachsrate neuer Nutzer. Weiterhin sind globale Daten wie die gesamte Anzahl aktueller Matches und die durchschnittliche Like-Quote



ablesbar. Während der Simulation können verschiedene Parameter wie die Art des Matching-Algorithmus justiert werden. Neben den simulationsbezogenen Attributen können auch Informationen des ausgewählten Nutzers betrachtet werden. Es kann zu jedem Nutzer durch einen Klick auf seinen Kreis angezeigt werden, wie viele andere Nutzer ihm bereits vorgeschlagen wurden, wie viele er gelikt hat und welche Matches er bereits hat. Auch Persönliche Informationen wie die Augenfarbe, Größe oder das Geschlecht des ausgewählten Nutzers können im Dashboard eingesehen werden.

## 6 Durchgeführte Experimente

Im folgenden Kapitel werden die verschiedenen Experimente und deren Auswirkung auf die Matching-Quote dargestellt. In der Simulationsstudie sollen vier verschiedene Szenarien untersucht werden, die sich aus den ex-ante festgelegten Laufzeitvariablen ergeben: 1) es wird ein *Preferred Proposing* verwendet, 2) die Berechnung eines Elo-Scores wird verwendet, 3) es werden Fakeprofile mit einbezogen und 4) es wird Werbung mit einbezogen. Als Benchmark für diese Szenarien wird ein Baseline-Szenario gebildet, indem keine der genannten Möglichkeiten verwendet wird. Den Agenten werden dort ausschließlich zufällige Profile vorgeschlagen. Anschließend wird analysiert, wie sich die Faktoren der Szenarien auf die Matching-Quote auswirken.

### 6.1 Erwartungen

Es wird erwartet, dass der Einbezug von der Berechnung des Elo-Scores und insbesondere eines *Preferred Proposing* zu einer erhöhten Matching-Quote führt. Der Elo-Score sorgt dafür, dass sich Personen mit gleicher Like-Wahrscheinlichkeit vorgeschlagen werden. Ein ähnlicher Elo-Score bedeutet eine ähnliche wahrgenommene Attraktivität der Nutzer. Es wird vermutet, dass sich Nutzer mit einer ähnlich wahrgenommenen Attraktivität auch eher Liken/Matchen. Weiterhin führt das *Preferred Proposing* dazu, dass dem Nutzer eher Profile mit dem bevorzugten Typ äußerlicher Kriterien vorgeschlagen werden und Personen, die einem ein Like gegeben haben. Ersteres sollte die Like-Quote drastisch erhöhen, wohingegen letzteres die Matching-Quote steigern sollte.

### 6.2 Numerische Ergebnisse

Das Baseline-Szenario dieser Simulation besteht aus 1.000 Nutzern und randomisierten Vorschlägen von möglichen Partnern. Es werden keine weiteren Laufzeitvariablen einbezogen. Anschließend werden mehrere Durchläufe zu dem *Preferred Proposing* durchgeführt und die Matching-Quote und *Expected Happiness* der gesamten Population verglichen. Weiterhin wird das *Preferred Proposing* um die Elo-Score Berechnung, um vorhandene Fakeprofile und Werbung erweitert. In *Tabelle 1* ist zu sehen ist, dass ein zufälliges Vorschlagen von anderen Nutzern zum einen zu einer sehr geringen Matching-Quote führt und zum anderen zusätzlich die Population unglücklich stimmt (die Population startet mit einer *Expected Happiness* von 50%).

	Matching-Quote	Expected Happiness
Random Proposing	0,7%	43%
Preferred Proposing	23,7%	63%
+ Elo-Score	26,8%	65%
+ Fakeprofil	27,3%	63,50%
+ Werbung	27,3%	59,50%

**Tabelle 1.** Ergebnis Experimente

Die Erwartungen aus 6.1 *Erwartungen* lassen sich sonst weitgehend bestätigen. Gemäß *Tabelle 1* hat das *Preferred Proposing* einen maßgeblichen Einfluss auf die Matching-Quote (Likes/Matches). Bei dem zufälligen Vorschlägen von Nutzern gibt es lediglich 0,7% Matches, wobei ein gezielter Algorithmus die Kennzahl auf fast ¼ erhöhen kann. Auch kann das wahrgenommene Glück der Population durch ein *Preferred Proposing* gesteigert werden. Der Einbezug des Elo-Scores kann die Matching-Quote weiter erhöhen. Da Fakeprofile alle anderen Nutzer liken, führt deren Einbezug zu einer weiteren Erhöhung der Matching-Quote. Sichtbar ist jedoch, dass das wahrgenommene Glück der Population um 1,5% nachlässt. Der Einbezug von Werbung hat keine Auswirkung auf die Matching-Quote, wirkt sich jedoch negativ auf die *Expected Happiness* der gesamten Population aus.

## 7 Diskussion und Fazit

Ziel dieser Arbeit war es das Verhalten einer heterogenen Population zu analysieren, visualisieren und einen Matching-Algorithmus zu entwerfen, der die Matching-Quote der Population maximiert. Dafür wurde im Rahmen der Simulationsstudie ein an *Tinder* angelehntes Modell einer Dating Plattform agentenbasiert konzeptioniert, implementiert und visualisiert. Bei der Konzeptionierung wurde darauf geachtet, dass das Verhalten der Agenten möglichst realitätsnah abgebildet wird. Die Agenten haben die Möglichkeit ein Swipe-Menü zu besuchen und andere Nutzer zu liken oder gematchten Nutzern eine Nachricht zu schreiben. Neben dem Nutzerverhalten wurde auch darauf Wert gelegt, die persönlichen Attribute und Motive jedes Nutzers möglichst realistisch darzustellen. Dafür wurde eine Abstraktion in vier Datingmotive vorgenommen und die persönlichen Attribute anhand von Quellen festgelegt. Anschließend wurde eine Implementierung und Visualisierung des konzeptionierten Modells in *AnyLogic PLE* vorgenommen. Anhand der implementierten Simulation war es möglich, das Nutzerverhalten weiter zu analysieren und Schlussfolgerungen für die Praxis zu ziehen.

Dazu wurden die Ergebnisse zufälliger Vorschläge und Vorschläge anhand eines erweiterten Matching-Algorithmus (*Preferred Proposing*) verglichen. Es konnte festgestellt werden, dass ein durchdachter Matching-Algorithmus maßgeblichen Einfluss auf den Erfolg einer Dating-Plattform hat. So kann die Matching-Quote von unter 1% auf knapp 25% erhöht werden. Darüber hinaus ist in dieser Arbeit die Erhöhung der Zufriedenheit der heterogenen Nutzerbasis sowie der Einfluss verschiedener Motive auf das Matching-Verhalten mit einbezogen worden. Weiterhin

wurde darauf geachtet eine für die Praxis nützliche Visualisierung der Methoden umzusetzen.

Erweiterbar ist die Simulation auf verschiedenen Dimensionen. Ein wichtiger Schritt wäre es, die Präferenzen und Profile von Nutzern weiter auszubauen. Mögliche Profilbeschreibung und deren Auswirkungen werden derzeit beispielsweise nicht betrachtet. Auch können Nutzer natürlich weitaus mehr Attribute und somit auch mehr Präferenzen als die Größe, Augen- und Haarfarbe besitzen. Denkbar wäre auch eine Erweiterung durch Machine-Learning-Techniken um anhand von vergangenen Ereignissen Vorhersagen darüber treffen zu können, ob ein Nutzer auf eine Nachricht antworten wird [9]. Weiterhin könnte die Simulation durch lernende Agenten erweitert werden. Weniger erfolgreiche Agenten könnten beispielsweise testen, ob sie durch die Verbindung ihres Social-Media-Profiles, durch bessere Bilder oder einen Beschreibungstext mehr Matches erhalten und somit im Laufe der Simulation ihr Profil anpassen. Interessant kann sich auch die Streuung von für den Nutzer attraktiven Profilen gestalten. So könnte es sein, dass der Nutzer insgesamt länger aktiv ist, wenn nur alle zehn Profile ein für ihn passendes vorgeschlagen wird und nicht alle zu Beginn. Dafür muss die Simulation jedoch auch um eine flexible online Zeit erweitert werden.

## Referenzen

1. Tong, S.T., Hancock, J.T., Slatcher, R.B.: Online dating system design and relational decision making: Choice, algorithms, and control. *Pers. Relatsh.* 23, 645–662 (2016).
2. Kleinerman, A., Rosenfeld, A., Ricci, F., Kraus, S.: Optimally balancing receiver and recommended users' importance in reciprocal recommender systems. In: *Proceedings of the 12th ACM Conference on Recommender Systems - RecSys '18*. pp. 131–139. ACM Press, New York, New York, USA (2018).
3. Singleboersen-Vergleich.de: DER ONLINE-DATING-MARKT 2017-2018. (2018).
4. Bitkom e.V.: Jeder dritte Deutsche sucht die Liebe im Internet, <https://www.bitkom.org/Presse/Presseinformation/Jeder-dritte-Deutsche-sucht-die-Liebe-im-Internet> (Accessed: 18.12.2019).
5. Krzywicki, A., Wobcke, W., Kim, Y.S., Cai, X., Bain, M., Mahidadia, A., Compton, P.: Collaborative Filtering for people-to-people recommendation in online dating: Data analysis and user trial. *Int. J. Hum. Comput. Stud.* 76, 50–66 (2015).
6. Krzywicki, A., Wobcke, W., Cai, X., Mahidadia, A., Bain, M., Compton, P., Kim, Y.S.: Interaction-Based Collaborative Filtering Methods for Recommendation in Online Dating. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 342–356 (2010).
7. Cai, X., Bain, M., Krzywicki, A., Wobcke, W., Kim, Y.S., Compton, P., Mahidadia, A.: Collaborative Filtering for People to People Recommendation in Social Networks. Presented at the (2010).
8. Terveen, L., McDonald, D.W.: Social matching: A Framework and Research Agenda. *ACM Trans. Comput. Interact.* 12, 401–434 (2005).
9. Xia, P., Jiang, H., Wang, X., Chen, C., Liu, B.: Predicting user replying behavior on a large online dating site. *Proc. 8th Int. Conf. Weblogs Soc. Media, ICWSM 2014*. 545–554 (2014).
10. Menkin, J.A., Robles, T.F., Wiley, J.F., Gonzaga, G.C.: Online dating across the life span: Users' relationship goals. *Psychol. Aging.* 30, 987–993 (2015).

11. Macal, C.M., North, M.J.: Tutorial on agent-based modelling and simulation. *J. Simul.* 4, 151–162 (2010).
12. Bonabeau, E.: Agent-based modeling: Methods and techniques for simulating human systems. In: *Proceedings of the national academy of sciences*. pp. 7280–7287 (2002).
13. Alsaleh, S., Nayak, R., Xu, Y., Chen, L.: Improving Matching Process in Social Network Using Implicit and Explicit User Information. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 313–320 (2011).
14. Qu, Y., Liu, H., Du, Y., Wu, Z.: Reciprocal Ranking: A Hybrid Ranking Algorithm for Reciprocal Recommendation. In: Geng, X. and Kang, B. (eds.) *PRICAI 2018: Trends in Artificial Intelligence*. pp. 455–463. Springer, Cham (2018).
15. Chen, L., Nayak, R., Xu, Y.: A Recommendation Method for Online Dating Networks Based on Social Relations and Demographic Information. In: *2011 International Conference on Advances in Social Networks Analysis and Mining*. pp. 407–411. IEEE (2011).
16. Pizzato, L.A., Rej, T., Yacef, K., Koprinska, I., Kay, J.: Finding Someone You Will Like and Who Won't Reject You. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 269–280 (2011).
17. Xia, P., Liu, B., Sun, Y., Chen, C.: Reciprocal Recommendation System for Online Dating. In: *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 - ASONAM '15*. pp. 234–241. ACM Press, New York, New York, USA (2015).
18. Wobcke, W., Krzywicki, A., Kim, Y.S., Cai, X., Bain, M., Compton, P., Mahidadia, A.: A Deployed People-to-People Recommender System in Online Dating. *AI Mag.* 36, 5 (2015).
19. Brozovsky, L., Petricek, V.: Recommender System for Online Dating Service. *arXiv Prepr. cs/0703042*. (2007).
20. Akehurst, J., Koprinska, I., Yacef, K., Pizzato, L., Kay, J., Rej, T.: CCR - A content-collaborative reciprocal recommender for online dating. *IJCAI Int. Jt. Conf. Artif. Intell.* 2199–2204 (2011).
21. Nayak, R.: Utilizing Past Relations and User Similarities in a Social Matching System. In: Huang, J.Z., Cao, L., and Srivastava, J. (eds.) *Advances in Knowledge Discovery and Data Mining*. pp. 99–110. Springer, Berlin, Heidelberg (2011).
22. Li, L., Li, T.: MEET: A generalized framework for reciprocal recommender systems. *ACM Int. Conf. Proceeding Ser.* 35–44 (2012).
23. Eirinaki, M., Vazirgiannis, M.: Web mining for web personalization. *ACM Trans. Internet Technol.* 3, 1–27 (2003).
24. McFee, B., Lanckriet, G.: Metric learning to rank. *ICML 2010 - Proceedings, 27th Int. Conf. Mach. Learn.* 775–782 (2010).
25. Law, A.: *Simulation modeling and analysis*. McGraw-Hill Education - Europe (2014).
26. Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F.: The ODD protocol: A review and first update. *Ecol. Modell.* 221, 2760–2768 (2010).
27. Pleines, C., Bothe, A., Schweyer, C.: *ZU-ZWEIT.de Pro lbild-Studie 2019 / 2020: Partnersuche, Strand und Gurken*. (2019).
28. AnyLogic: AnyLogic Simulation Software, <https://www.anylogic.de/use-of-simulation/multimethod-modeling/> (Accessed: 28.02.2019).