

Die Rolle der mehrschichtigen Netzwerkanalyse im Bereich Social Media Analysis am Beispiel der Identifizierung von Persönlichkeitsmerkmalen

Kai Heinrich¹, Adrian Fischer¹, Michael Seifert¹

¹Technische Universität Dresden

{kai.heinrich, adrian.fischer, michael.seifert3}@tu-dresden.de

Abstract. Die Netzwerktheorie erfährt gerade in den letzten Jahren einen Aufschwung: Während ursprünglich mit Hilfe klassischer Graphen lediglich gleichartige Beziehungen zwischen gleichartigen Knoten dargestellt werden konnten, besteht durch die Verfügbarkeit von großen Mengen an Netzwerkdaten, insbesondere aus dem Bereich der sozialen Medien, die Notwendigkeit komplexere Netzwerkmodelle einzusetzen. Diese sind bspw. in der Lage, unterschiedliche Kantentypen und Knotenarten zu berücksichtigen. In diesem Beitrag wird das bisher eindimensional betrachtete Problem der Identifikation von Persönlichkeitsmerkmalen anhand von Social-Media-Daten betrachtet. Dabei existieren im Hinblick auf die Analyse der Persönlichkeit zurzeit kein geeignetes mehrschichtiges Netzwerkmodell sowie keine Systematisierung von geeigneten Kennzahlen zur Auswertung eines solchen Modells. Zur Bearbeitung dieser Forschungslücken wird zunächst ein passendes mehrschichtiges Netzwerkmodell unter Einbeziehung von Facebook-Daten erstellt und anschließend passende Kennzahlen zur Auswertung vorgeschlagen. Diese Kennzahlen werden dann in einer Korrelationsstudie mit empirisch erhobenen Ausprägungswerten der Persönlichkeitsmerkmale verglichen. Als Ergebnis können Korrelationen zwischen mehreren Strukturmerkmalen der Netzwerke und den Persönlichkeitsmerkmalen nachgewiesen werden. Abschließend werden daraus verschiedene Hypothesen zum Umgang mit sozialen Medien unter Berücksichtigung der jeweiligen Persönlichkeitsmerkmale aufgestellt.

Keywords: Multilayer Network Analysis, Personality Traits, Big Five, Social Media Analysis, Facebook

1 Einleitung

Bei der Untersuchung sozialer Netzwerke galt bisher das Erlangen von zuverlässigen, umfangreichen und auswertbaren Datenbeständen als eine der größten Herausforderungen, so dass Kompromisse hinsichtlich der Tiefe und des Umfangs von Experimenten eingegangen werden mussten [1, 2]. Im Zuge der Digitalisierung können elektronische Interaktionen, wie sie in Social-Media-Plattformen stattfinden, aufgezeichnet,

¹⁵th International Conference on Wirtschaftsinformatik,
March 08-11, 2020, Potsdam, Germany

in großen Mengen gespeichert sowie zu geringen Kosten abgefragt und für Forschungszwecke genutzt werden [3, 4]. Diese Daten werden unter anderem benutzt, um digitale Spuren mit Hilfe von Netzwerkanalysen zu untersuchen, um damit Rückschlüsse auf Verhaltensmuster und Persönlichkeitsmerkmale von Nutzern sozialer Netzwerke zu ziehen [5]. Die Darstellung und Analyse von Netzwerken mithilfe klassischer Verfahren der Netzwerkanalyse erfolgt indem Nutzer als Knoten durch Kanten miteinander in Beziehung gesetzt werden [6]. Diese klassische Form der Netzwerktheorie limitiert allerdings die Möglichkeiten der Auswertung, was angesichts der Menge und Heterogenität der verfügbaren Daten zu einem sub-optimalen Prozess der Wissensgenerierung führt [7, 8]. Die Erweiterung der klassischen Theorie um beliebige Knoten- sowie Beziehungsarten führt zum Konzept der mehrschichtigen Netzwerke (Multilayer Network - MLN) und deren Analyse (MLNA). Dies ermöglicht beispielsweise nicht nur Freundschafts- sondern auch Misstrauensbeziehungen zu analysieren. Während zwar eine fundierte Theorie zu MLN existiert, sind die Anwendungspotenziale bzgl. der Analyse weitestgehend unbeachtet. Weiterhin ist die Eignung einzelner Kennzahlen aus dem Bereich MLNA für den Bereich der Analyse sozialer Medien (Social Media Analysis - SMA) noch nicht festgestellt. Die Ziele dieses Beitrages ergeben sich aus dieser Forschungslücke: Die Rolle der MLNA im Bereich SMA soll durch eine Korrelationsstudie zur Identifikation des Zusammenhangs zwischen multidimensionalen Netzwerk-kennzahlen und Persönlichkeitsmerkmalen aufgezeigt werden. Der Beitrag zur Erreichung dieser Forschungsziele ist wie folgt gegliedert: In Abschnitt 2 werden grundlegende Konzepte der mehrschichtigen Netzwerkanalyse und der Analyse von Persönlichkeitsmerkmalen vorgestellt. Abschnitt 3 beschäftigt sich mit der Beschreibung der Studie und des angewandten Netzwerkmodells samt Kennzahlen. In Abschnitt 4 werden die Ergebnisse der Korrelationsstudie vorgestellt. In Abschnitt 5 werden die Ergebnisse diskutiert sowie eine sowie ein Ausblick auf zukünftige Forschungsperspektiven gegeben.

2 Konzeptueller Hintergrund

2.1 Grundlagen mehrschichtiger Netzwerke

Um komplexere Sachverhalte darstellen zu können wird der klassische Graph erweitert [8, 9]. Ein Beispiel für einen MLN-Graph (b) im Vergleich mit einem klassischen Graph (a) ist in Abbildung 1 gegeben.

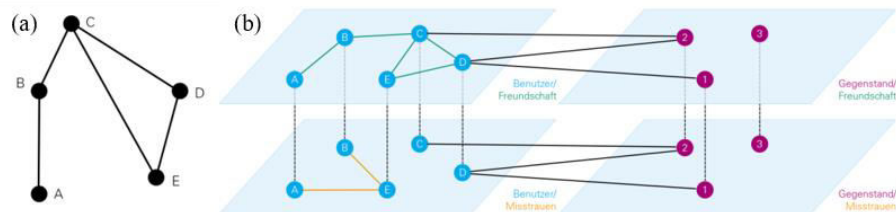


Abbildung 1. Vergleich klassischer Graph (a) mit mehrschichtigem Netzwerk (b)

Zusätzliche Ebenen stellen verschiedene Arten von Beziehungen (z. B. Freundschaft und Misstrauen) zwischen gleichen Knoten oder Verbindungen zwischen unterschiedlichen Arten von Knoten (z. B. Personen und Gegenstände) dar. Soll beides dargestellt werden, müssen neben Ebenen auch Aspekte, welche die Unterscheidungsmerkmale (Knotentyp und Verbindungsart) zwischen den Ebenen beinhalten, eingeführt werden. Diese lassen sich auch als Dimensionen verstehen. Da ein Knoten in mehreren Ebenen sein kann, muss zwischen Knoten und Knoten-Ebenen-Tupeln unterscheiden werden. Nur zwischen letzterem können maximal eine ungerichtete Kante bzw. zwei gerichtete Kanten bestehen. Es existiert nun die Möglichkeit zwischen Intra- und Inter-Ebenen-Kanten zu unterscheiden. Ein Sonderfall von Inter-Ebenen-Kanten sind die Kopplungskanten, welche denselben Knoten über Ebenen verbindet. Insbesondere sei noch darauf verwiesen, dass mit den einzelnen Ebenen auch eine zeitliche Komponente dargestellt werden kann.

MLN können verschiedenartig dargestellt bzw. kategorisiert werden. Sie lassen sich in vier Typen unterscheiden. Zuerst in *Multiplex-Netzwerke* und *knotenetikettierte Netzwerke*, welche nach verschiedenen Kantenarten bei gleichbleibender Bedeutung der Knoten differenzieren bzw. unterschiedliche Knotentypen über verschiedene Ebene abbilden. Diese beiden Typen kommen vergleichsweise häufig in der Literatur vor und besitzen als Sonderfall eine Zeiterweiterung. Selten treten dagegen *kombinierte Netzwerkformen* und *Hypernetzwerke* auf. Kombinierte Netzwerkformen vereinen die beiden Unterscheidungsmerkmale Verbindungsarten und Knotentypen. Hypernetzwerke, bei denen Kanten über mehr als zwei Knoten bestehen können, lassen sich als Erweiterung von klassischen Graphen verstehen.

Bei Multiplex-Netzwerken sind Aspekte und Ebenen bezüglich unterschiedlicher Kantenarten bei gleicher Knotenart zulässig. Dabei muss ein Knoten nicht in allen Ebenen existent sein, d. h. das Netzwerk ist nicht knotenbeständig. Gleichzeitig ergeben Inter-Ebenen-Kanten zwischen unterschiedlichen Knoten-Ebenen-Tupeln keinen Sinn, da jeder Knoten in jeder Ebene vorhanden sein kann. Multiplex-Netzwerken sind immer diagonal, da alle Inter-Ebenen-Kanten nur Kopplungskanten sind, also zwischen demselben Knoten innerhalb verschiedener Ebenen. Gleichzeitig muss ein solches Netzwerk nicht kategorisch sein, d. h. zwischen allen Entitäten eines Knotens über diverse Ebenen muss keine Verbindung bestehen (siehe [10–14]).

Unter knotenetikettierten Netzwerken werden alle Netzwerke verstanden, bei denen ein einziges Etikett jedem Knoten zugeordnet ist. Jedes Etikett kann als Ebene aufgefasst werden, wenn dieses Netz in die generelle Form mehrschichtiger Netzwerke umgewandelt wird. Diese Netzwerke sind Ebenen-disjunkt, da jeder Knoten in nur einer Ebene liegt.

2.2 Kennzahlen zur Auswertung mehrschichtiger Netzwerke

Die Kennzahlen zur Auswertung eines generierten oder empirisch aufgestellten MLN im Bereich SMA lassen sich anhand der Facette des Netzwerkes, zu der sie eine

Aussage treffen, systematisieren. Neben bekannten, knotenspezifischen Kennzahlen gibt weitere Kategorien von Kennzahlen bzgl. Kanten, Ebenen oder dem gesamten Netzwerk. Zur Analyse der Persönlichkeit kommen ausschließlich knotenspezifische Kennzahlen und in Frage, da wir hier die Persönlichkeit eines einzelnen Subjekts untersuchen bzw. die Relevanz einer einzelnen Ebene (z.B. Freundschaft) für die jeweiligen Persönlichkeitsmerkmale. Weitere Kategorien von Kennzahlen spielen zunächst keine Rolle, da diese sehr aggregierten Charakter haben und daher keine Rückschlüsse mehr auf die Eigenschaften eines Subjekts schließen lassen.

Tabelle 1. Systematisierte Übersicht von knotenspezifischen Kennzahlen für die MLNA

| Kennzahl | Unter-kategorie | Netzwerk-voraussetzung | Anfor-derungen | Quelle |
|--------------------------------|-----------------|------------------------|----------------|-------------------------|
| Globale Nachbarn | EGO | MPN | - | Hristova et al. 2016 |
| Eingangsnachbarn | EGO | MPN | G | Berlingerio et al. 2011 |
| Ausgangsnachbarn | EGO | MPN | G | |
| Kern-Nachbarn | EGO | MPN | - | Hristova et al. 2016 |
| Kern-Eingangsnachbarn | EGO | MPN | G | |
| Kern-Aus.-Nachbarn | EGO | MPN | G | |
| Kern(<i>n</i>)-Nachbarn | EGO | MPN | - | |
| Kern(<i>n</i>)-Ein.-Nachbarn | EGO | MPN | G | |
| Kern(<i>n</i>)-A.-Nachbarn | EGO | MPN | G | |
| Reziprozität | EGO | MLN | G | |
| Relative Reziprozität | EGO | MLN | G | |
| Kantenzahl | EGO | MLN | - | Suo et al. 2015 |
| Kantenzahl (streng) | EGO | MLN | - | |
| Eingangskantenzahl | EGO | MLN | G | |
| Ausgangskantenzahl | EGO | MLN | G | |
| Gewichtsb. Kantenzahl | EGO | MLN | W | |
| Gew. Ein.-Kantenzahl | EGO | MLN | G,W | |
| Gew. Aus.-Kantenzahl | EGO | MLN | G,W | |
| Berühmtheit (konst.) | GLO | MPN | G,W | Khadangi & Bagheri 2017 |
| Berühmtheit (nutzerb.) | GLO | MPN | G,W | |
| Berühmtheit (global) | GLO | MPN | G,W | |
| Unbeliebtheit (konst.) | GLO | MPN | G,W | |
| Unbeliebtheit (nutzerb.) | GLO | MPN | G,W | |
| Unbeliebtheit (global) | GLO | MPN | G,W | |

Eine Übersicht der knotenspezifischen Kennzahlen ist in Tabelle 1 dargestellt. Dabei dienen die fettgedruckten Abkürzungen im Text als Referenz zum Tabellenkopf und

die fett und kursiv gedruckten Bezeichnungen als Referenz hinsichtlich der möglichen Ausprägungen dieser Merkmale.

Die Kennzahlen lassen sich dabei nach **Unterkategorien** klassifizieren. *Egozentrische Kennzahlen (EGO)* benötigen nur das Ego-Netzwerk, also sämtliche Knoten, die direkt mit diesem verbunden sind, sowie alle zugehörige Kanten (also auch Kanten zwischen den Nachbarknoten untereinander). *Globale Kennzahlen (GLO)* ermitteln Werte, welche auf dem gesamten Netzwerk, also auch indirekte Verbindungen der mit dem Ego-netzwerk verbundenen Knoten. Neben globalen und egozentrischen Kennzahlen, lassen sich die Kennzahlen weiter bzgl. der Einsatzvoraussetzungen hinsichtlich des benötigten Netzwerktyps (**Netzwerkvoraussetzung**) und besonderen Eigenschaften wie gerichtete (*G*) und/oder gewichtete Kanten (*W*) oder die Zulässigkeit von Inter- und/oder Intra-Ebenen-Kanten (**Anforderungen**) unterscheiden. Dabei wird bzgl. der Netzwerkvoraussetzung unterschieden zwischen Kennzahlen, welche explizit ein Multiplex Netzwerk fordern (*MPN*) und Kennzahlen, welche für beliebige MLN angewandt werden können (*MLN*).

Anhand der wiss. Literatur können folgende knotenspezifische Kennzahlengruppen unterschieden werden:

Nachbarschaft. Die globale Nachbarschaft leitet sich von dem eindimensionalen Konzept des Knotengrades ab und gibt die Menge der Knoten an mit dem der fokussierte Knoten über mindestens eine Ebene und mindestens eine Kantenart verbunden ist. Bei gerichteten Graphen werden hier zusätzlich zwischen eingehenden und ausgehenden Kanten die Kennzahlen Eingangs- und Ausgangsnachbarn unterschieden. Im Gegensatz zur globalen Nachbarschaft gibt die Kernnachbarschaft nur die Anzahl der verbunden Knoten an, mit denen der fokussierte Knoten über alle Ebenen oder alle bestehenden Kantenarten verbunden ist. Eine Abschwächung kann hier durch die Berechnung der n-Varianten berücksichtigt werden, in denen ein Knoten nur über n Ebenen verbunden sein muss [15, 16].

Kantenzahl. Da in Hypergraphen zwei Knoten durch mehr als eine Kante verbunden sein können gibt die Kantenzahl die Menge aller Kanten (auch verschiedenen Typs) an, welche mit einem Knoten verbunden sind. Die strenge Variante der Kantenzahl berücksichtigt weiterhin nur eine Verbindung zwischen zwei Knoten. Auch hier kann in gerichteten Graphen wieder zwischen Eingangs- und Ausgangskantenzahl unterschieden werden. Die gewichteten Varianten der Kantenzahl berücksichtigen Kantengewichte, so dass verschiedene Verbindungen eine Stärke der Verbindung zugeordnet werden kann, anstatt alle Verbindungen als gleichgewichtet zu betrachten. Dies setzt voraus, dass Kantengewichte definiert sind [17].

Reziprozität. Die Reziprozität gibt für einen Knoten in einem gerichteten Graphen an, wie viele von den ausgehenden Knoten auch „erwidert“ werden. Teilt man die Anzahl der reziproken Knoten durch die Gesamtanzahl der anliegenden Knoten erhält man die relative Reziprozität [11].

Berühmtheit/Unbeliebtheit. Neben der Abstrahierung von Kennzahlen klassischer Graphen wurden auch gänzlich neue Instrumente für Multiplex-Netzwerke geschaffen. Dazu zählen die Kennzahlen Berühmtheit und Unbeliebtheit. Zur Ermittlung der Kennzahlen wird zunächst die Differenz aus Eingangs- und Ausgangsgewicht eines Knotens für jede Ebene einzeln berechnet und diese mit einem knotenspezifischen Reinheitswert multipliziert. Für die Berühmtheit wird die Differenz in der Reihenfolge *Eingangsgewicht- Ausgangsgewicht* gebildet und bei der Unbeliebtheit genau anders herum. Für die Berechnung des Reinheitswerts kommen entweder die Entropie oder der Gini-Index eines Knotens in Frage, die sich beide ebenfalls aus den knoten- und ebenenspezifischen Eingangs- und Ausgangsgewichten berechnen lassen. Anschließend wird über alle Ebenen hinweg ein gewichteter Mittelwert gebildet, so dass spezielle Ebenen in der Berechnung deutlicher hervorgehoben oder abgeschwächt werden können. Bei der konstanten Variante werden alle Ebenen gleich gewichtet. Bei der nutzerbasierten Variante wird eine invertierte Ebenenrelevanz für den das jeweilige Ego-Netzwerk benutzt, während bei der globalen Variante die invertierte durchschnittliche Ebenenrelevanz aller betrachteten Ego-Netzwerke herangezogen wird [13]. Wie bei der Reziprozität ist hier die Voraussetzung, dass es sich um einen gerichteten Graphen mit vorhandenen Kantengewichten handelt [18].

Ebenerelevanz. Die Ebenerelevanz eines Knotens bzgl. einer Ebene ist streng genommen eine knotenspezifische Kennzahl, welche aber für jede Ebene einzeln angegeben werden muss. Sie gibt den Anteil aller Nachbarn eines Knotens für eine Ebene an. Die XOR-Variante gibt den Anteil an welcher ausschließlich über eine spezielle Ebene verbunden ist. Ein Mittelweg zwischen den beiden Konzepten stellt die gewichtete Ebenerelevanz dar: Hier wird für jeden Nachbarn in der Ebene ermittelt, über wie viele andere Ebenen er noch erreichbar wäre.

2.3 Identifizierung von Persönlichkeitsmerkmalen

Ein erklärtes Ziel des Zusammenwirkens der Persönlichkeitspsychologie und der Informatik/Wirtschaftsinformatik lautet: Computer sollen auf Basis der digitalen Verhaltensdaten eines Individuums (z. B. aus sozialen Medien) dessen Persönlichkeit präzise bestimmen können [19, 20]. Der Nutzen entsteht hierbei durch das mit Persönlichkeitsmerkmalen assoziierte Verhalten, welche eine Beurteilung einer Person im Hinblick auf verschiedene Faktoren zulässt. Als Beispiele können hier Beurteilung im Personalwesen oder des Kaufverhaltens und die damit einhergehende Verbesserung von Personen- bzw. Produktempfehlungssystemen angeführt werden. Als Grundlage zur Bestimmung der Persönlichkeit gibt es verschiedene verhaltenspsychologische Theorien und Paradigmen, die allerdings eine eher gespaltene wissenschaftliche Basis bilden. Im Mittelpunkt der Diskussion steht dabei die Dissonanz zwischen biologischen und gesellschaftlichen Einflüssen auf unsere Persönlichkeit. Die Untersuchung der letzteren Einflussgruppe wird auch unter dem Eigenschaftsparadigma subsumiert und in diesem Beitrag aufgrund ihrer Nähe zur

Netzwerktheorie weiterverfolgt. Dabei steht insbesondere die Auswertung von personen-zentrierten Netzwerken im Vordergrund (Ego-Netzwerke). Eine ausgedehnte Diskussion der verschiedenen Paradigmen würde an dieser Stelle zu weit führen und es wird auf [21] verwiesen. Zur Quantifizierung der Persönlichkeit hat sich vor allem im Bereich der SMA das Modell der „Big Five“ etabliert. In Anlehnung an Goldberg können dabei die Persönlichkeitsmerkmale *Offenheit (OH)*, *Gewissenhaftigkeit (GWH)*, *Extraversion (EX)*, *Verträglichkeit (VER)* sowie *Neurotizismus (NR)* identifiziert werden. Zur Analyse dieser Merkmale im Bereich SMA können zwei Gruppen von wiss. Beiträgen unterschieden werden: rein assoziative Studien, welche das Ziel haben die Merkmale möglichst präzise vorherzusagen (Beispiele: [22, 23]) und korrelative Studien, welche das Ziel haben den eigentlichen Zusammenhang zu erklären (Beispiele: [24, 25]). Während es einige Ansätze der MLNA in der ersten Gruppe zum Einsatz kommen (siehe [22, 26]), wurde das Multi-Schichten-Netzwerkmodell im Bereich der korrelativen Studien bisher nicht beachtet. Im Hinblick auf Forschungsziel (c) und der komplexen Struktur der Persönlichkeitsmerkmale wird hier daher eine korrelative Studie mithilfe des MLNA-Ansatzes durchgeführt.

3 Design der Studie

3.1 Datengrundlage und Vorgehen

Eine initiale Gruppe an Probanden vom Umfang $n=30$ wurde mithilfe einer Anreizmechanik gewonnen. Aus persönlichkeitspsychologischer Sicht ist eine Referenzgruppe zu benennen, die den Probanden im Alter gleichen muss [19]. Hier wurde daher eine Gruppe zwischen (einschließlich) 23 und 27 Jahren betrachtet. Aufgrund des geringen Stichprobenumfangs wurde die Gruppe nicht weiter geteilt und alle Probanden einer Referenzgruppe zugewiesen.

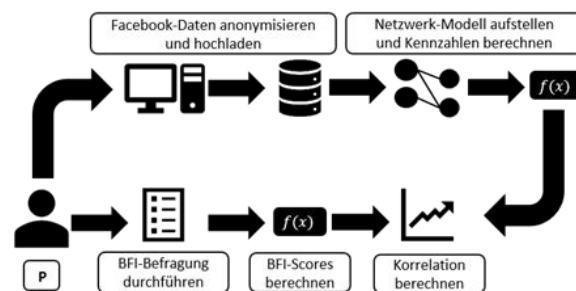


Abbildung 2. Vorgehensmodell zur Korrelationsstudie

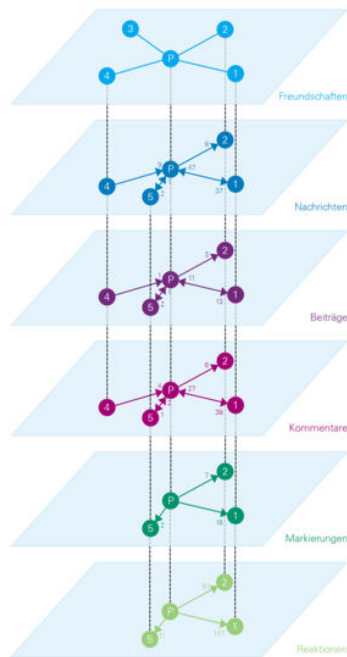
Um eine Korrelation zu bestimmen, wurden einerseits eine Vielzahl an Facebook-Daten der Probanden erhoben (z.B.: Freundschaften, Nachrichten, Markierungen) und auf der anderen Seite eine Befragung zur Bestimmung der Persönlichkeitsmerkmale durchgeführt. Der Forschungsprozess ist in Abbildung 2 dargestellt. Die Facebook-

Daten wurden mithilfe eines bereitgestellten Tools bereits anonymisiert hochgeladen, sodass lediglich eine ID-Nummer den Fragebogen und die Daten verbindet. Zusätzlich wurde auf die Erhebung von Textinhalten der Nachrichten und Posts zum Schutz der Privatsphäre der Probanden verzichtet.

Zur Erhebung der Merkmalsausprägungen wurde das Big Five Inventory (BFI) in deutscher Fassung verwendet und anschließend die BFI-Scores für alle Persönlichkeitsmerkmale der Probanden berechnet [27, 28].

3.2 Netzwerkmodell und Kennzahlen

Aus den vorliegenden Facebook-Daten wird ein Multi-Plex-Netzwerk konstruiert. Dabei werden die Probanden stets als P dargestellt.



(a) Multiplex-Netzwerk

| Komplexe (multiplexe) Kennzahl | Berücksichtigte Ebenen | | | | | |
|--------------------------------|------------------------|-------------|----------|------------|------|------------|
| | Freundschaft | Nachrichten | Beiträge | Kommentare | Tags | Reaktionen |
| Globale Nachbarn | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Eingangsnachbarn | | ✓ | ✓ | ✓ | | |
| Ausgangsnachbarn | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kern-Nachbarn | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kern-Eingangsnachbarn | | ✓ | ✓ | ✓ | | |
| Kern-Aus-Nachbarn | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kern(n)-Nachbarn | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kern(n)-Ein.-Nachbarn | | ✓ | ✓ | ✓ | | |
| Kern(n)-A.-Nachbarn | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Reziprozität | | ✓ | ✓ | ✓ | | |
| Relative Reziprozität | | ✓ | ✓ | ✓ | | |
| Kantenzahl | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kantenzahl (streng) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Eingangskantenzahl | | ✓ | ✓ | ✓ | | |
| Ausgangskantenzahl | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gewichtsb. Kantenzahl | | ✓ | ✓ | ✓ | | |
| Gew. Ein.-Kantenzahl | | ✓ | ✓ | ✓ | | |
| Gew. Aus.-Kantenzahl | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Berühmtheit (konst.) | | ✓ | ✓ | ✓ | | |
| Berühmtheit (nutzerb.) | | ✓ | ✓ | ✓ | | |
| Berühmtheit (global) | | ✓ | ✓ | ✓ | | |
| Unbeliebtheit (konst.) | | ✓ | ✓ | ✓ | | |
| Unbeliebtheit (nutzerb.) | | ✓ | ✓ | ✓ | | |
| Unbeliebtheit (global) | | ✓ | ✓ | ✓ | | |

(b) Verwendete multiplexe Kennzahlen

Abbildung 3. Netzwerk-Modell und verwendete multiplexe Kennzahlen

Die erste Ebene charakterisiert das Freundschaftsnetzwerk und beinhaltet alle Freundschaftsbeziehungen zu P. Da es hier lediglich eine binäre Option gibt, sind die Kanten nicht gewichtet. Alle weiteren Ebenen werden als gerichtete und gewichtete Aktivitätsgraphen dargestellt. Die Ebene Nachrichten gibt an, wer an wen wie viele Nachrichten verschickt hat. Analog dazu werden die Ebenen Kommentare und

Beiträge definiert, indem ausgehende Verbindungen die Beiträge bzw. Kommentare von P ggü. Dritten berücksichtigt und eingehende Verbindungen Beiträge bzw. Kommentare Dritter ggü. P. Die weiteren Aktivitätsgraphen umfassen nur ausgehende Daten der Nutzer, da für die andere Richtung keine Daten vorliegen. So zeigt die Ebene Markierungen, wie häufig P von anderen markiert wurde. Die Ebene Reaktion umfasst sämtliche Reaktionen (z. B. „Like“) von P ggü. anderen. Als Grundlage des Netzwerkmodells wird ein Multiplex-Netzwerk gewählt, da hier Aspekte und Ebenen bezüglich unterschiedlicher Kantenarten bei gleicher Knotenart zulässig sind (siehe Abschnitt 2.1). Dies ermöglicht den größtmöglichen Spielraum zur Aufstellung verschiedener Netzwerk Kennzahlen, da die meisten Kennzahlen im Bereich der MLN für Multiplex-Netzwerke definiert sind. Abbildung 3 (a) zeigt die beispielhafte Darstellung eines Multi-Plex Ego-Netzwerkes mit allen genannten Ebenen für einen Probanden.

Gemäß dem Vorgehen aus Abschnitt 4.1 müssen nach Erfassung der Daten und nach

Aufstellung des Netzwerkmodells durch Beschreibung seiner Ebenen, die Kennzahlen zur Auswertung des Modells definiert werden. Abbildung 3 (b) zeigt die benutzten mehrdimensionalen Kennzahlen zur Auswertung mit Referenz auf die in Abschnitt 4.1 definierten Netzwerkebenen.

4 Auswertung der Ergebnisse

Tabelle 2 zeigt die deskriptive Auswertung des BFI mit Referenzwerten nach [29]. Sofern man die Daten nach Lang et al. 2001 aufgrund der höheren Teilnehmerzahl als „repräsentativer“ erachtet, waren die Pretest-Probanden überdurchschnittlich extravertiert (+0,27 Punkte) und verträglich (+0,33 Punkte). Der Mittelwert der Gewissenhaftigkeit der Pretest-Probanden weicht kaum vom Mittelwert nach Lang et al. 2001 ab (+0,02 Punkte). Dagegen gaben die Probanden im Vergleich an, unterdurchschnittlich neurotisch (-0,36 Punkte) sowie offen (-0,14 Punkte) zu sein.

Tabelle 2. Netzwerk-Modell und verwendete multiplexe Kennzahlen

| Persönlichkeitseigenschaft | Pretest-Probanden | | Lang et al. 2001 (Junge Erwachsene) | |
|----------------------------|-------------------|-----------------|--|-----------------|
| | Mittelwert | Std.-Abweichung | Mittelwert | Std.-Abweichung |
| <i>Extraversion</i> | 3,80 | 0,73 | 3,53 | 0,68 |
| <i>Verträglichkeit</i> | 3,83 | 0,50 | 3,50 | 0,52 |
| <i>Gewissenhaftigkeit</i> | 3,64 | 0,63 | 3,62 | 0,52 |
| <i>Neurotizismus</i> | 2,29 | 0,76 | 2,65 | 0,60 |
| <i>Offenheit</i> | 3,41 | 0,62 | 3,55 | 0,52 |

Tabelle 3 zeigt die Ergebnisse der Korrelationsanalyse für die multiplexen Kennzahlen. Auffällig sind die vielen hochsignifikanten Korrelationen zwischen Extraversion und verschiedenen Kennziffern. Extravertiertere Probanden scheinen demnach auf das Gesamtnetzwerk gesehen mehr Interaktionspartner zu haben als weniger extravertierte Personen. Dieser Trend ist auch bei den verschiedenen

Versionen der Kern-Nachbarn-Kennzahlen zu beobachten. Dies bedeutet, dass Probanden mit höheren Extraversionen nicht nur mit mehr Dritten interagierten, sondern dass diese Interaktion auch in erhöhtem Maße auf mehreren Ebenen gleichzeitig stattgefunden hat. Es besteht dabei nicht nur eine signifikante Rangkorrelation zwischen Extraversion und der Anzahl der Interaktionspartner, sondern auch zwischen Extraversion und der Häufigkeit der Interaktion (gewichtsbasierte Kantenzahlen).

Tabelle 3. Ergebnisse der Korrelationsanalyse

| | Extra- version | Verträg- lichkeit | Gewissen- haftigkeit | Neuroti- zismus | Offen- heit |
|---|---------------------------|------------------------------|---------------------------------|----------------------------|------------------------|
| Globale Nachbarn | 0,335* | 0,212 | 0,157 | -0,075 | 0,194 |
| (Globale) Eingangsnachbarn | 0,469** | 0,252 | 0,117 | 0,017 | 0,230 |
| (Globale) Ausgangsnachbarn | 0,364** | 0,222 | 0,019 | -0,076 | 0,110 |
| Kern(2)-Nachbarn | 0,452** | 0,250 | 0,115 | -0,024 | 0,185 |
| Kern(2)-Eingangsnachbarn | 0,370** | 0,103 | 0,072 | -0,055 | 0,116 |
| Kern(2)-Ausgangsnachbarn | 0,403** | 0,248 | 0,045 | -0,012 | 0,080 |
| Kern(3)-Nachbarn | 0,404** | 0,277* | 0,067 | -0,021 | 0,061 |
| Kern-Eingangsnachbarn | 0,224 | 0,050 | -0,042 | -0,094 | 0,083 |
| Kern(3)-Ausgangsnachbarn | 0,352** | 0,153 | 0,007 | 0,024 | 0,019 |
| Kern(4)-Nachbarn | 0,356** | 0,148 | 0,002 | 0,029 | 0,024 |
| Kern(4)-Ausgangsnachbarn | 0,188 | 0,010 | 0,018 | -0,013 | 0,085 |
| Kern(5)-Nachbarn | 0,191 | 0,013 | 0,016 | -0,016 | 0,088 |
| Kern-Ausgangsnachbarn | 0,198 | 0,200 | 0,049 | -0,081 | 0,161 |
| Kern-Nachbarn | 0,198 | 0,200 | 0,049 | -0,081 | 0,161 |
| Reziprozität | 0,495** | 0,217 | 0,096 | -0,005 | 0,272* |
| Relative Reziprozität | -0,045 | -0,038 | 0,100 | 0,207 | 0,176 |
| Kantenzahl | 0,419** | 0,222 | 0,153 | -0,028 | 0,213 |
| Kantenzahl (streng) | 0,405** | 0,212 | 0,162 | -0,028 | 0,223 |
| Eingangskantenzahl | 0,478** | 0,252 | 0,112 | 0,012 | 0,239 |
| Ausgangskantenzahl | 0,368** | 0,189 | 0,014 | -0,075 | 0,115 |
| Gewichtsbasierte Kantenzahl | 0,405** | 0,108 | 0,014 | -0,071 | 0,269* |
| Gew. Eingangskantenzahl | 0,368** | 0,085 | -0,010 | -0,061 | 0,302* |
| Gew. Ausgangskantenzahl | 0,363** | 0,085 | 0,043 | -0,047 | 0,246 |
| Berühmtheit (konstant) | 0,089 | -0,120 | -0,088 | -0,290* | 0,253 |
| Berühmtheit (nutzerbasiert) | 0,174 | 0,026 | -0,036 | -0,186 | 0,202 |
| Berühmtheit (global) | 0,164 | -0,031 | -0,026 | -0,139 | 0,258* |
| Unbeliebtheit (konstant) | 0,192 | 0,130 | -0,069 | -0,031 | -0,091 |
| Unbeliebtheit (nutzerbasiert) | 0,171 | 0,042 | -0,124 | -0,052 | 0,030 |
| Unbeliebtheit (global) | 0,150 | 0,045 | -0,131 | -0,116 | -0,106 |
| Fettdruck Die Korrelation ist auf dem 0,10 Niveau signifikant (zweiseitig). | | | | | |
| * Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig). | | | | | |
| ** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). | | | | | |

Interessant ist, dass diese Zusammenhänge sowohl für ein- als auch ausgehende Aktivitäten gefunden wurden. Die verstärkte Interaktion von extravertieren Personen scheint also nicht einseitig, sondern bidirektional gewesen zu sein. Weiterhin sind die Ergebnisse hinsichtlich der verschiedenen Varianten der Reziprozität bemerkenswert.

Die (absolute) Reziprozität weist eine signifikante Korrelation mit der Extraversion auf. Teilt man die (absolute) Reziprozität aber durch die Anzahl der Ausgangsnachbarn (relative Reziprozität), so wurde ein solcher Zusammenhang nicht

gefunden. Dies ist ein Hinweis dafür, dass extravertiertere Personen verstärkt die Interaktionen mit Dritten initiierten. Gemessen an der Anzahl von Nachbarn, zu denen eine ausgehende Interaktion besteht, verzeichnen sie nämlich keine erhöhte Anzahl eingehender Interaktionen. Absolut verfügen sie aber über mehr Interaktionspartner. Dies erklärt wiederum, warum keine signifikanten, positiven Korrelationen zu den verschiedenen Versionen der Berühmtheit bestehen. Ähnliche Beobachtungen wie bei der Extraversion lassen sich auch bei der Verträglichkeit tätigen, allerdings in deutlich geringerem Umfang. Hinsichtlich der Offenheit besteht auf höherem Signifikanzniveau mit Eingangskantenzahl und globale Eingangsnachbarn eine signifikante Korrelation. Die signifikante Korrelation mit der Reziprozität sowie die fehlenden signifikanten Korrelationen zu ausgehenden Kennzahlen deuten darauf hin, dass Zusammenhänge zu ungerichteten Kennzahlen im Wesentlichen durch die entsprechenden eingehenden Kennzahlen verursacht werden. Auf offenere Probanden schien also vermehrt zugegangen worden zu sein. Dieser Zusammenhang wird je nach Signifikanzniveau von ein oder zwei Versionen der Berühmtheit bestätigt. Weiterhin nennenswert ist die signifikante negative Rangkorrelation zwischen Neurotizismus und zumindest der konstanten Version der Berühmtheit.

Dies könnte ein Indikator dafür sein kann, dass andere Nutzer die Interaktion mit neurotischen Personen als schwieriger empfanden. Weitere Zusammenhänge zwischen Neurotizismus und globalen Netzwerkkennzahlen wurden allerdings nicht gefunden. Genauso wurden keine signifikanten Korrelationen mit der Gewissenhaftigkeit entdeckt. Tabelle 4 zeigt zum Vergleich die an den gleichen Daten durchgeführte Analyse auf Basis eindimensionaler Kennzahl für die Ebene der Freundschaftsbeziehungen¹. Hier wird lediglich ein Teil der oben dargestellten Zusammenhänge bestätigt, insbesondere die signifikante Korrelation zwischen Offenheit, Gewissenhaftigkeit und Anzahl der Freunde sind hierbei zu nennen. Vergleicht man die ganzheitliche Betrachtung durch MLNA mit einer rein ebenspezifischen Betrachtung, so erhält man durch die Betrachtung aller Ebenen stabilere und weitere signifikante Korrelationen.

Tabelle 4. Ergebnisse der Korrelationsanalyse für die Ebene Freundschaft

| Kennzahl | Extra- version | Verträg- lichkeit | Gewissen- haftigkeit | Neuro- tizismus | Offenheit |
|---|---------------------------|------------------------------|---------------------------------|----------------------------|------------------|
| Knotengrad | 0,237 | 0,165 | 0,186 | -0,052 | 0,283* |
| Nachbarn _{XOR} | 0,045 | 0,094 | 0,281* | -0,094 | 0,312* |
| Ebenenrelevanz | -0,157 | -0,127 | 0,134 | 0,066 | 0,040 |
| Ebenenrelevanz _{XOR} | -0,283* | -0,236 | 0,115 | 0,024 | 0,035 |
| Gewichtete Ebenenrelevanz | -0,232 | -0,222 | 0,129 | 0,066 | 0,068 |
| Fettdruck Die Korrelation ist auf dem 0,10 Niveau signifikant (zweiseitig). * Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig). ** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). | | | | | |

¹ Die vollständigen Ergebnisse bzgl. aller Ebenen sowie die Daten sind als Supplement-File anonymisiert einsehbar unter: <https://gofile.io/?c=iXtjKB>

Die durchgeführte Korrelationsstudie ermöglicht eine Ableitung von Hypothesen, um den Zusammenhang zwischen den Aktivitäten in sozialen Medien und der verschiedenen Persönlichkeitseigenschaften zu verdeutlichen. Die abgeleiteten Hypothesen für die verschiedenen Persönlichkeitsmerkmale sind in Tabelle 5 dargestellt.

Tabelle 5. Abgeleitete Hypothesen bezüglich der verschiedenen Persönlichkeitsmerkmale

| | Formulierung | Merkmal |
|------|---|----------------|
| H1.1 | Personen mit hohen Extraversionswerten verfügen in sozialen Medien über vergleichsweise viele Interaktionspartner, sowohl auf das Gesamtnetzwerk gesehen, als auch in den einzelnen Interaktionsebenen (z. B. Nachrichten- oder Kommentar-Netzwerke). | <i>EX</i> |
| H1.2 | Personen mit hohen Extraversionswerten kommunizieren mit überverhältnismäßig vielen ihrer Interaktionspartner auf mehreren Interaktionsebenen der sozialen Netzwerke gleichzeitig. | <i>EX</i> |
| H1.3 | Personen mit hohen Extraversionswerten kommunizieren in sozialen Medien vergleichsweise häufig über ihre Freundschaften-Netzwerke hinaus. | <i>EX</i> |
| H1.4 | Personen mit hohen Extraversionswerten weisen nicht nur relativ viele Interaktionspartner in sozialen Medien auf, sondern kommunizieren mit diesen auch vergleichsweise häufig. | <i>EX</i> |
| H1.5 | Personen mit hohen Extraversionswerten haben in sozialen Medien auffällig weniger Scheu, Inhalte öffentlich zu teilen. | <i>EX</i> |
| H1.6 | In sozialen Medien ist die Kommunikation zwischen Personen mit hohen Extraversionswerten und Dritten bidirektional, d. h. es werden auch überverhältnismäßig viele Antworten verzeichnet. Dabei sind extravertiertere Personen vergleichsweise häufig diejenigen, die die Interaktion initiieren. | <i>EX</i> |
| H1.7 | Die verstärkte Interaktion von Personen mit hohen Extraversionswerten in sozialen Medien beschränkt sich primär auf eigene Themen. Mit ausgehenden Betrachtungen in Sub-Graphen, die Interesse an fremden Themen bekunden (z. B. Kommentar-Netzwerken), werden daher keine positiven Zusammenhänge erwartet. | <i>EX</i> |
| H2.1 | Für Personen mit hohen Verträglichkeitswerten spielen in sozialen Medien Private-Nachrichten-Netzwerke gemessen an der Häufigkeit der Interaktion eine besonders geringe Rolle | <i>VER</i> |
| H2.2 | Personen mit hohen Verträglichkeitswerten hinterlassen in sozialen Medien besonders häufig Beiträge in Chroniken anderer Nutzer, die auch in den eigenen Chroniken Beiträge verfassen. Dabei ist unklar, wer diese „Kettenreaktion“ initiiert. | <i>VER</i> |
| H2.3 | Personen mit hohen Verträglichkeitswerten nutzen in sozialen Medien vergleichsweise häufig simple Reaktionsmöglichkeiten (z. B. „Gefällt mir“-Angaben) auf die Inhalte dritter Nutzer. Dabei kommt solchen Reaktionen-Netzwerken bei verträglicheren Nutzern im Hinblick auf deren gesamte Kommunikation eine besonders große Rolle zu. | <i>VER</i> |
| H3.1 | Personen mit hohen Gewissenhaftigkeitswerten wählen in sozialen Medien ihre Kommunikationspartner mit Bedacht aus. Dadurch besitzen sie viele Kontakte in ihren Freundschaften-Netzwerken, mit denen sie online gar nicht interagieren | <i>GWH</i> |
| H3.2 | Personen mit hohen Gewissenhaftigkeitswerten kommentieren in sozialen Medien besonders ungern Beiträge anderer Nutzer. | <i>GWH</i> |
| H4 | Personen mit hohen Neurotizismus-Werten sind in sozialen Medien „unattraktiver“ Kommunikationspartner, was besonders durch niedrigere Werte in den Beliebtheits-Kennzahlen belegt wird. | <i>NR</i> |
| H5.1 | Personen mit hohen Offenheitswerten sind in sozialen Medien begehrtere Interaktionspartner. Auf sie gehen vergleichsweise viele Nutzer zu und mit ihnen wird besonders häufig interagiert. Dies schlägt sich auch in den Berühmtheits-Kennzahlen nieder. | <i>OH</i> |
| H5.2 | Personen mit hohen Offenheitswerten haben besonders viele Kontakte in den Freundschaften-Netzwerken sozialer Medien, von denen sie dann aber mit einer | <i>OH</i> |

| | | |
|------|---|----|
| | relativ geringen Zahl interagieren. | |
| H5.3 | Personen mit hohen Offenheitswerten kommunizieren am liebsten über Private-Nachrichten-Netzwerke sozialer Medien, wodurch diesen gemessen an ihrer Kommunikationshäufigkeit insgesamt eine übergeordnete Rolle zukommt. | OH |
| H5.4 | Personen mit hohen Offenheitswerten nutzen Reaktionen-Netzwerke sozialer Medien in geringerem Umfang. | OH |

5 Zusammenfassung & Fazit

Durch die durchgeführte Studie wurden teilweise signifikante Korrelation zwischen MLNA-Kennzahlen und Persönlichkeitsmerkmalen auf alleiniger Basis der Strukturinformation der Ego-Netzwerke festgestellt. Dabei stellte sich insbesondere heraus, dass einige Zusammenhänge über mehrere Ebenen stabil waren während andere nur bei bestimmten Ebenen und damit bei bestimmten Social-Media-Daten auftraten. Damit stellt sich der Vorteil der MLNA ggü. der Betrachtung isolierter, eindimensionaler Ego-Netzwerke heraus. Insbesondere zwischen den Netzwerkkennzahlen und der Extraversion, aber auch der Offenheit konnten signifikante Rangkorrelationen bestätigt werden.

Dies deckt sich mit den Ergebnissen von Meta-Analysen im Forschungsfeld [30, 31], nach welchen sich gerade die beiden soeben genannten Persönlichkeitseigenschaften mit Daten aus sozialen Medien besonders gut vorhersagen lassen und sich dieser Effekt für die Extraversion noch einmal verstärkte, sobald Netzwerkkennzahlen als Eingangsvariablen verwendet werden. Diese Ergebnisse zeigen, dass MLNA zur Identifikation von einigen Persönlichkeitsmerkmalen verwendet werden kann. Insbesondere zwischen die hier identifizierten Merkmale, wie Extrovertiertheit, können beispielsweise Rückschlüsse auf kognitive Fähigkeiten, Wahrscheinlichkeit des Gemütszustands sowie Zusammenhänge mit dem Kaufverhalten hergestellt werden [32, 33]. Die Aussagen der Studie unterliegen allerdings einigen Limitationen. Zu diesen zählen der geringe Stichprobenumfang sowie die aus datenschutzrechtlichen Gründen nicht einbezogenen inhaltlichen Daten. Weiterhin wurde die Entwicklung der Persönlichkeit im untersuchten Alterssegment als konstant angesehen und nicht beachtet. Aufgrund der Komplexität der menschlichen Persönlichkeit ist die Anwendung von simplen Netzwerkanalysen alleine nicht ausreichend, um die Persönlichkeit vollständig zu beschreiben [34]. Die Hypothesen können nur deskriptiver Natur sein und stellen den Ausgangspunkt weiterer induktiver Forschungsarbeiten dar. Doch bereits die vorläufigen Ergebnisse liefern ein tiefergehendes Verständnis, wie mithilfe von Daten aus sozialen Medien selbstbeurteilte Persönlichkeitseigenschaften von Computern vorhergesagt bzw. berechnet werden könnten. Dabei ergibt sich aus der Persönlichkeitsdefinition im Rahmen dieses Beitrags eine weitere, „philosophische“ Limitation: Die Autoren verstehen unter den selbstbeurteilten „Big Five“- Persönlichkeitsfaktoren ein nützliches Instrument zur ersten, groben Einschätzung bisher unbekannter Menschen, dass nicht der vollen Komplexität der menschlichen Persönlichkeit gerecht wird. Dies bedeutet, dass selbst wenn auf Grundlage der Ergebnisse dieser Arbeit Computer

eines Tages selbstbeurteilte Persönlichkeitseigenschaften besser abschätzen können als Menschen, diese damit nicht auch zwangsweise eine bessere Einschätzung der tatsächlichen, vollumfänglichen Persönlichkeit vornehmen können. Die hier vorgestellte Analyse ist damit lediglich als erstes grobes Instrument zur Einschätzung einer Persönlichkeit zu verstehen und ist stets als Ergänzung zu anderen Methoden, welche auf lexikalischen oder verhaltenstechnischen Studien basieren, zu verstehen [35].

Literatur

1. Rogers, E.M.: Progress, problems and prospects for network research. *Social Networks*. 9, 285–310 (1987).
2. Venturini, T., Latour, B.: *The Social Fabric: Digital Traces and Quali-quantitative Methods*. Presented at the Proceedings of Future En Seine 2009: The Digital Future of the City (2010).
3. Whelan, E., Teigland, R., Vaast, E., Butler, B.: Expanding the horizons of digital social networks. *Information and Organization*. 26, 1–12 (2016).
4. Natarajan, M.: Role of Text Mining in Information Extraction and Information Management. *DESIDOC Bulletin of Information Technology*. 25, 31–38 (2005).
5. Hansen, D.L., Shneiderman, B., Smith, M.A.: Introduction to Social Media and Social Networks. In: Hansen, D.L., Shneiderman, B., and Smith, M.A. (eds.) *Analyzing Social Media Networks with NodeXL. Insights from a Connected World*. pp. 3–9. Morgan Kaufmann Publishers, Burlington, USA (2010).
6. Wang, D., Pedreschi, D., Song, C., Giannotti, F., Barabasi, A.-L.: Human mobility, social ties, and link prediction. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. pp. 1100–1108. ACM (2011).
7. Barabási, A.-L., Albert, R., Jeong, H.: Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*. 272, 173–187 (1999).
8. Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A.: Multilayer networks. *Journal of Complex Networks*. 2, 203–271 (2014).
9. Bollobás, B. ed: *Modern Graph Theory*. Springer New York, New York, NY (1998).
10. Szell, M., Lambiotte, R., Thurner, S.: Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences of the United States of America*. 107, 13636–13641 (2010).
11. Filiposka, S., Gajduk, A., Dimitrova, T., Kocarev, L.: Bridging online and offline social networks. *Physica A: Statistical Mechanics and its Applications*. 471, 825–836 (2017).
12. Berlingerio, M., Coscia, M., Giannotti, F., Monreale, A., Pedreschi, D.: The pursuit of hubbiness. *Journal of Computational Science*. 2, 223–237 (2011).
13. Bindu, P.V., Thilagam, P.S., Ahuja, D.: Discovering suspicious behavior in multilayer social networks. *Computers in Human Behavior*. 73, 568–582 (2017).
14. Sharma, R., Magnani, M., Montesi, D.: Effects of missing data in multilayer networks. *Soc. Netw. Anal. Min.* 6, 509 (2016).

15. Hristova, D., Noulas, A., Brown, C., Musolesi, M., Mascolo, C.: A multilayer approach to multiplexity and link prediction in online geo-social networks. *EPJ Data Sci.* 5, 203 (2016).
16. Socievole, A., Rango, F.D., Caputo, A.: Opportunistic mobile social networks. *Computer Communications.* 87, 1–18 (2016).
17. Suo, Q., Sun, S., Hajli, N., Love, P.E.D.: User ratings analysis in social networks through a hypernetwork method. *Expert Systems with Applications.* 42, 7317–7325 (2015).
18. Khadangi, E., Bagheri, A.: Presenting novel application-based centrality measures for finding important users based on their activities and social behavior. *Computers in Human Behavior.* 73, 64–79 (2017).
19. Soldz, S., Vaillant, G.E.: The Big Five Personality Traits and the Life Course. *Journal of Research in Personality.* 33, 208–232 (1999).
20. Ozer, D.J., Benet-Martínez, V.: Personality and the Prediction of Consequential Outcomes. *Annual Review of Psychology.* 57, 401–421 (2006).
21. Neyer, F.J., Asendorpf, J.B.: Personality–Relationship Transaction in Young Adulthood. *Journal of Personality and Social Psychology.* 81, 1190–1204 (2001).
22. Farnadi, G., Sitaraman, G., Sushmita, S., Celli, F., Kosinski, M., Stillwell, D., Davalos, S., Moens, M.-F., Cock, M.: Computational Personality Recognition in Social Media. *User Model User-Adap Inter.* 26, 109–142 (2016).
23. Kleanthous, S., Herodotou, C., Samaras, G., Germanakos, P.: Detecting Personality Traces in Users' Social Activity. Presented at the Social Computing and Social Media. 8th International Conference, Toronto, ON, Canada (2016).
24. Liu, D., Campbell, W.K.: The Big Five Personality Traits, Big Two Metatraits and Social Media. *Journal of Research in Personality.* 70, 229–240 (2017).
25. Farnadi, G., Zoghbi, S., Moens, M.-F., Cock, M.: Recognising Personality Traits Using Facebook Status Updates. *AAAI Technical Report.* (2013).
26. Bai, S., Hao, B., Li, A., Yuan, S., Gao, R., Zhu, T.: Predicting Big Five Personality Traits of Microblog Users. Presented at the Proceedings of the 2013 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Workshops WI-IAT 2013, Atlanta, GA, USA (2013).
27. Rammstedt, B.: Die deutsche Version des Big Five Inventory (BFI), (1997).
28. Rammstedt, B., Danner, D.: Die Facettenstruktur des Big Five Inventory (BFI). *Diagnostica.* 63, 70–84 (2017).
29. Lang, F.R., Lüdtke, O., Asendorpf, J.B.: Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica (Online-Ausgabe).* 47, 111–121 (2001).
30. Azucar, D., Marengo, D., Settanni, M.: Predicting the Big 5 Personality Traits from Digital Footprints on Social Media. *Personality and Individual Differences.* 124, 150–159 (2018).
31. Kaushal, V., Patwardhan, M.: Emerging Trends in Personality Identification Using Online Social Networks—A Literature Survey. *ACM Trans. Knowl. Discov. Data.* 12, 1–30 (2018).
32. Feiler, D.C., Kleinbaum, A.M.: Popularity, Similarity, and the Network Extraversion Bias. *Psychological Science.* 26, 593–603 (2015).

33. McNiel, J.M., Fleeson, W.: The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality*. 40, 529–550 (2006).
34. Youyou, W., Kosinski, M., Stillwell, D.: Computer-based Personality Judgments Are More Accurate than Those Made by Humans. *Proceedings of the National Academy of Sciences of the United States of America*. 112, 1036–1040 (2015).
35. Bleidorn, W., Hopwood, C.J.: Using Machine Learning to Advance Personality Assessment and Theory. *Personality and Social Psychology Review*.