

Die automatische Identifizierung von AfD Anhänger auf Twitter: Der Vanellus-Follower-Classification-Index VaFCI.

von

Dr. Christian „Chrilly“ Donninger

Nimzowerkstatt KG, A-3925 Altmelon 110

c.donninger@wavenet.at

<http://www.godotfinance.com>

Zusammenfassung:

Vanellus (der Kiebitz) ist ein von Dr. Christian „Chrilly“ Donninger entwickelter Twitter-Bot. Es werden rund um die Uhr alle deutschsprachigen Tweets herunter geladen und in einer Datenbank gespeichert. Diese Datenbank kann nach vielfältigen Kriterien durchsucht und analysiert werden. Dieser Artikel beschreibt die automatische Identifizierung von AfD Anhängern mit Hilfe des Vanellus Follower Classification Index VaFCI. Der VaFCI ist nicht auf diese Fragestellung beschränkt. Es werden die Anhänger der AfD für diese Untersuchung verwendet, weil die AfD auf Twitter präsent ist und die User polarisiert. Es ist für einen Menschen anhand der Tweets eines Users relativ einfach zu entscheiden, ob er Pro oder Contra eingestellt ist. Es schlüpfen dem VaFCI ein paar offensichtliche AfD-Anhänger durch. Wenn er jedoch „*anschlägt*“, dann ist die Klassifizierung fast immer richtig.

Die Daten:

Der Vanellus-Bot lädt seit dem 27. April 2018 rund um die Uhr in Echtzeit alle deutschsprachigen Tweets mit dem Suchkriterium „lang:de“ vom Twitter-Server und speichert diese in einer eigens entwickelten, sehr effizienten Datenbank. Die Anzahl der Tweets beträgt rund 1 Million/Tag. Für diese Untersuchung wurden die im Juni 2018 gesendeten 29,3 Millionen Tweets heran gezogen. Es waren in diesem Zeitraum 3,5 Millionen User aktiv.

AfDKompakt	Mitgliedermagazin AfD
AfDimBundestag	Bundestags Fraktion AfD
AfD	Offizieller Account AfD
AfDwirkt	AfDwirkt AfD
AfD_HD	Offizieller Account Heidelberg AfD
StBrandner	Stephan Brandner AfD
PetrBystronAfD	Petr Bystron AfD
JoanaCotar	Joana Cotar AfD
GottfriedCurio	Dr. Gottfried Curio AfD
Frank_Pasemann	Frank Pasemann AfD
Beatrix_vStorch	Beatrix von Storch AfD
Alice_Weidel	Alice Weidel AfD
BjoernHoecke	Björn Höcke AfD Thüringen
Joerg_Meuthen	Prof. Dr. Jörg Meuthen MdEP AfD
MalteKaufmann	Dr. Malte Kaufmann AfD
SteinbachErika	Erika Steinbach parteilos

Es wurden für die 16 Accounts in der Tabelle die Follower herunter geladen. Die ersten 5 User sind allgemeine AfD-Accounts. Es folgen die persönlichen Accounts der - auf Twitter - 10 wichtigsten AfD-Politiker. Zusätzlich noch die ehemalige CDU-Politikerin Erika Steinbach. Die Liste dieser Accounts wird im folgenden AfD-16 genannt. Erika Steinbach ist kein AfD-Mitglied, sie wird jedoch dem AfD Umfeld zugerechnet. Sie hat mit 83.454 mehr Follower als alle anderen AfD-Politiker. Zusätzlich spricht sie einen etwas anderen Personenkreis an. Fügt man zu den ersten 15 Accounts einen weiteren AfD-Politiker hinzu, so erhöht sich die Anzahl der Follower fast nicht mehr. Die meisten folgen bereits mindestens einem der 15 zuvor Aufgelisteten. Bei

Erika Steinbach sind es hingegen 43.000 zusätzliche Follower. Dies ist für diese Untersuchung – und vermutlich auch für die AfD – ein interessanter Effekt.

Die Follower wurden am 25. Juli 2018 herunter geladen. In Summe haben die AfD-16 Accounts 464.479 Follower, zieht man Mehrfachnennungen ab, so kommt man auf eine Menge von 232.330 Usern. Man kann einen User nur klassifizieren, wenn man Information über ihn in Form von Tweets hat. Es werden alle User klassifiziert, die im Laufe des Juni 2018 mindestens 10 Tweets versandt haben. Das sind 17.630 User. Für diese Gruppe wird der VaFCI berechnet. Ein höherer Wert bedeutet,

der User identifiziert sich stärker mit der AfD. Für die Klassifikation werden die folgenden Kriterien heran gezogen:

Retweets:

Ein Retweet ist ein Tweet der Form „RT @username“. Wobei der @username in der Menge von AfD-16 vorkommen muss. Ein Retweet drückt üblicher Weise eine Zustimmung aus. Es Retweeten vereinzelt auch explizite Gegner Tweets von AfD-Politikern. Dies kommt vor, wenn sich nach Meinung des Gegners der ursprüngliche Tweet „von selbst richtet“. Dies betrifft primär Funktionäre auf Ortsgruppen Ebene. Bei den führenden AfD-Politikern kann man diesen Effekt ignorieren.

Hashtags:

Ein Hashtag ist ein Wort mit einem vorangestellten „#“. Man könnte – so wie dies bei der Klassifizierung von Hunderrassen durch die FCI (Fédération Cynologique Internationale) gemacht wird – eine Liste von typischen Merkmalen bzw. Hashtags von AfD Anhängern aufstellen. Das ist erstens mühsam und zweitens sind Hashtags nicht statisch. Man hechelt der Entwicklung von neuen Themen hinterher.

Beim Hunde-FCI halten sich die Züchter hingegen weitgehend an die vorgegebenen Merkmale. Diese sind allerdings wieder an als Prototypen angesehenen Hunden orientiert. Das Bild links zeigt einen



Ausschnitt aus Ferdinands Waldmüllers Gemälde „Niederösterreichische Bauernhochzeit“ (1843). Es ist auf zahlreichen Gemälden von Waldmüller dieser Typus von Bauernhund zu sehen. Dieser Typus war das Vorbild für den Rassestandard des Österreichischen Pinschers. Das Bild rechts zeigt Emil vom Hochmoor, einen heutigen Österreichischen Pinscher.



Die FCI-Klassifikation funktioniert, weil Rassestandards von Experten interpretiert werden. Eine automatische Klassifikation auf Grund dieses Regelwerkes würde absurde Fehlklassifikationen ergeben.

Beim VaFCI werden die relevanten Hashtags hingegen aus den Tweets der AfD-16 Accounts gebildet. Es ist jemand ein AfD-Anhänger, wenn er die von Alice Weidel&Co aktuell verwendeten Hashtags in seinen Tweets übernimmt. Die Groß- und Kleinschreibung wird dabei ignoriert.

Diese Methode hat einen Haken. Es verwenden auch explizite Gegner Hashtags wie „#afd“. Dieser – und andere – Hashtag sagt nur aus, dass sich ein User mit dem Thema AfD auseinandersetzt. Er hat hingegen keine Aussagekraft über die Einstellung zur AfD. Wenn man den Hashtag „#afd“ zur Klassifizierung verwendet, würden auch ausgesprochene Gegner einen hohen AfD-VaFCI aufweisen.

ContraAfD1	ContraAfD1	Die Lösung des Problems ist: Es werden zusätzlich die Tweets von drei sehr aktiven und prononcierten AfD Gegnern heran gezogen. Beim User @ContraAfD1 geht die Ausrichtung bereits aus dem Usernamen eindeutig hervor.
Chat Atkins	Klaus Jarchow	
DerDude72445273	Der Dude	

Diese drei Accounts – im folgenden ContraAfD-3 - waren dem Autor aus früheren Untersuchungen bekannt. Es wird – genauso wie bei den AfD-16 Accounts – für diese Gruppe die Menge der relevanten Hashtags ermittelt. Ein von ContraAfD-3 verwendeter Hashtag wird für die Klassifizierung ignoriert. Wie bereits erwähnt ist „#afd“ ein Beispiel dafür. Es wird jedoch ein Gegner nicht den Hashtag „#merkelmussweg“ verwenden.

In mathematischen Worten: Die Menge der relevanten Hashtags ist die Menge der relevanten AfD-16 minus der Menge der relevanten ContraAfD-3 Hashtags. Ein Hashtag ist relevant, wenn er mindestens vier Mal verwendet wurde (fünf- oder sechs-Mal ergibt praktisch dasselbe Ergebnis).

Die Schindel zur Lösung des AfD-Frankenthal Problems:

Ein offensichtlicher Test für einen Index ist: Man überprüft, ob User die sich im Usernamen eindeutig als Pro-AfD deklarieren, als solche erkannt werden. Legt man der Klassifikation nur die Retweets und die Hashtags zu Grunde, dann hat der User *@AfD_Frankenthal* einen Wert von 0.

Presseschau: Asylbetrug in Berlin: Hauptverdächtiger hatte 14 Identitäten <https://t.co/dmcVxFFUnl>

AfD Kompakt: Messer-Angriffe verbreiten Angst und Schrecken <https://t.co/wR9ANQ22dw> <https://t.co/WVQQfA3c8q>

AfD Kompakt: „Kundenfreundlichste Behörde Deutschlands“ <https://t.co/ELJTegWW5O> <https://t.co/rqgJpRKClX>

Presseschau: IC-Messerstecher soll auch seinen Nachbarn bedroht haben <https://t.co/Lfe5DzZbkG>

AfD Kompakt: Weniger „Europa“ ist mehr und besser für alle <https://t.co/AqnCQS9QUw> <https://t.co/HYAEIklogP>

AfD Kompakt: Körperverletzungen in Leipzig zwischen 2010 und 2017 um 350 Prozent gestiegen <https://t.co/KsLu1tS2VP> <https://t.co/1kfcY11G8t>

Wie man in diesem Ausschnitt sieht, gibt der für Twitter zuständige Funktionär aus Frankenthal die Meinung von AfD-Kompakt direkt wieder. Er sucht sich aus Pressemitteilungen AfD-kompatible Tweets heraus. Er verwendet jedoch nicht die üblichen Twitter-Regeln. Es gibt in den Tweets der AfD-Frankenthal keinen einzigen als solchen gekennzeichneten Retweet und keinen Hashtag. Man muss daher eine zusätzliche Klassifikationsmethode verwenden, die nicht die für Twitter typischen Konventionen voraus setzt.

Der [Shingling \(Dt. Schindel\) Algorithmus](#) ist eine 1993 erfundene Methode zur allgemeinen Klassifizierung von Dokumenten. Man möchte z.B. wissen, ob ein Artikel in die Kategorie Sport, Wetter, Politik ... fällt. Man gibt Muster-Berichte vor. Ein neuer Artikel wird entsprechend der größten Übereinstimmung klassifiziert. Die Muster-Berichte sind in diesem Fall die Tweets von AfD-16. Die erste Schindel besteht aus den Zeichen 1..11 des Textes. Die zweite Schindel von 2..12, die dritte von 3..13. Die Textfragmente überlappen sich wie die Schindeln eines Daches. Ein üblicher Wert ist eine Schindellänge von 9. Für diese Analyse hat der Wert 11 etwas besser funktioniert. Man berechnet wie bei den Hashtags aus den Tweets von AfD-16 die Menge der Schindel. Ebenso aus den Tweets von ContraAfD-3 die Menge der gegnerischen Schindel und zieht diese von den AfD-16 Schindel ab. Ein User ist AfD affin, wenn er viele für die AfD-typischen Schindel in seinen Tweets verwendet. Eine handverlesene Menge von Schindel ist praktisch unmöglich, man kann sie nur dynamisch aus den Tweets ermitteln. Es ist unvermeidlich, dass auch ein Gegner einzelne AfD-Schindel verwendet. Man muss eine Toleranz einbauen und erst über diesen Wert zu zählen beginnen. Diese allgemeine Methode erweist sich – nicht nur bei der AfD-Frankenthal – als erstaunlich treffsicher.

Reply:

Ein Reply ist ein Tweet der Form „@username“ zu Beginn des Textes (siehe [1] und [2]). Wobei @username in der Menge der AfD-16 Accounts vorkommen muss. Ein Reply drückt eher – aber nicht notwendiger Weise – eine Kritik aus. Dies gilt vor allem für User, die keine Retweets verwenden.

Die VaFCI-Formel:

Es wird jeweils für Retweets, Hashtags und Schindel die Ratio per Tweet ermittelt. Wenn jemand 10 mal retweeted, dann macht es einen Unterschied aus, ob er 20 oder 100 Tweets gesendet hat. Im ersten Fall ist die Retweet-Ratio 0,5 im zweiten 0,1. Bei den Replies wird für die Ratio durch die Anzahl der Tweets*(Retweets+1) dividiert. Wenn ein User keinen Retweet gesendet hat, dann ist dieser Wert die gewöhnliche Ratio. Wenn er sowohl Retweets als auch Replies verwendet, dann wird die Ratio entsprechend der Anzahl der Retweets verkleinert.

$$\text{VaFCI} = 100 * \text{Retweets} - 100 * \text{Replies} + 10 * \max(\text{Hashtags} - 0.2, 0) + 10 * \max(\text{Schindel} - 5.0, 0.0)$$

$\max(\text{Hashtags} - 0.2, 0.0)$ bedeutet: Es wird von der Hashtag-Ratio 0.2 abgezogen, weil auch Gegner manchmal AfD-Hashtags verwenden. Ist dieser Wert jedoch kleiner als 0, dann wird 0 und nicht diese

negative Zahl verwendet. Dasselbe gilt für die Schindel-Ratio. Durch drehen an diesen Toleranzwerten kann man die Sensibilität der Methode kalibrieren. Im Moment ist sie so eingestellt, dass nur definitive AfD-Anhänger erkannt werden. Verkleinert man die Schwellenwerte, dann schlägt der Index eher an, dafür klassifiziert man fälschlicher Weise User als AfD-Anhänger ein, die keine sind. Das ist ein bekanntes Problem bei der Kalibrierung von Alarmanlagen. Replies werden wie bereits erwähnt als Zeichen von politischer Ablehnung interpretiert. Das Gewicht verkleinert sich mit der Anzahl der Retweets.

Was ist ein User?

Bei der Klassifikation von Usern ergibt sich die Frage, was überhaupt ein User ist. Üblicher Weise verbindet man einen User mit einem bestimmten Usernamen wie z.B. *@Alice_Weidel*. Allerdings ist das nicht das interne, von Twitter, verwendete Konzept. Man erhält bei der Anmeldung eine – 8-Byte große – Zahl zugewiesen. Die interne Kennung von *@Alice_Weidel* ist 833398209053605888. Diese Kennung bleibt erhalten, wenn Alice Weidel ihren Usernamen z.B. in *@AliceWeidel* umbenennen würde. Lädt man von einem bestimmten User die Follower herunter, dann erhält man eine Liste von internen Kennungen und nicht die Namen. Man kann allerdings mit dieser Kennung die zu diesem User gespeicherten Informationen abfragen und würde bei der Nummer 833398209053605888 u.A. *@Alice_Weidel* erhalten. Diese Information nennt man Metadata.

Nachdem für die meisten Benutzer der Username die relevante Information ist, wurde für diese Untersuchung die Klassifikation nach Namen und nicht nach interner Kennung vorgenommen. Es wird jedoch im Output gekennzeichnet, wenn jemand seinen Namen geändert hat und so in der Auswertung mehrmals vorkommt. Wenn jemand seinen Namen ständig ändert, dann ist das – nach dem Motto „er wird schon einen Grund dafür haben“ – eine durchaus interessante Information. Der Output ist nach der internen Kennung aufsteigend sortiert. Man kann für weitere Untersuchungen einen User über diese Kennung identifizieren.

Das extremste Beispiel ist der User mit der Kennung 1367260729. Er hat im Juni 2018 unter 7

1367260729	Warumfrager	182	5	2	7	2069	66
1367260729	Weltenbummler01*	514	17	5	145	6910	89
1367260729	Kreuztraeger **	118	2	0	8	1751	100
1367260729	JohnDoe_2018***	73	2	0	12	1033	94
1367260729	Rosenkav alier02****	1143	31	5	94	13851	74
1367260729	Fragezeichen01*****	75	6	0	10	1370	141
1367260729	Querdenker007*****	963	32	4	125	14451	103

verschiedenen Namen getweetet. Das Zeichen ,*‘ am Ende ist nicht Teil des Usernamens. Es dient zur Kennzeichnung von Namenswechsel. Die erste Spalte nach dem Namen gibt die Anzahl der Tweets wieder. User 1367260729 hat im Juni unter verschiedenen Identitäten 3068 Tweets versendet. Die letzte Spalte ist der AfD-VaFCI. Ein Wert über 60 bedeutet, der User ist mit sehr hoher Wahrscheinlichkeit ein AfD-Anhänger. 1367260729 gehört – auch wenn man sich seine Tweets durchliest – zur „*Rechtspopulistischen-Hardcore-Community*“. Er versendet im Durchschnitt 100 Tweets pro Tag. Die 4. Spalte gibt die Anzahl der Retweets wieder. Diese Zahl bezieht sich nur auf die Retweets von AfD-16 Accounts. Nummer 1367260729 retweetet wesentlich mehr Tweets aus seiner Community. Dies ist ein eigener Diskurs, der in seiner Radikalität wesentlich über jene der AfD-16 Politiker hinaus geht. Es werden immer wieder User aus dieser Community von Twitter gesperrt. Die Namensänderung ist ein ungeeignetes Mittel, solchen Sperrungen zu entkommen. Wenn Vanellus weiß, dass User „*Warumfrager*“ identisch mit dem „*Weltbummler01*“ ... ist, dann weiß dies natürlich auch Twitter.

Die Top-10:

Die Grafik unten zeigt die Top-10 Follower gemessen an der Anzahl der Tweets. Der mit großen Abstand führende @ffd365 ist ein Nachrichten-Bot. Er versendet pro Tag rund 2.000 Tweets, die er über diverse Nachrichtenkanäle bezieht. Er gibt auch AfD-Meldungen weiter. User @ffd365 wird richtiger Weise mit einem AfD-Index von 0 bewertet. Er ist ein politisch weitgehend neutraler Verteiler von Nachrichten.

Die Nummer 3, @mrstone0656 ist – was die Anzahl der Tweets angeht – eine zentrale Figur der Rechtspopulistischen-Hardcore Community. Dasselbe gilt für die Plätze 4 bis 10.

Follower-Id	Follower	Tweets ▼	Retweets	Replies	Hashtags	Shingles	VaFCI
98190279	ffd365	60233	0	0	30	196192	0
479445216	FWWinterberg	10818	0	0	33	43475	0
816717166900809729	mrstone0856	10774	643	4	945	196720	139
745264589458247680	Renft1964	10329	364	0	710	148789	98
3457564343	e_pitzky	10011	489	3	761	176069	131
887758764232847360	DeutschZuerst	9501	668	0	838	193935	161
947517423409942528	OnymOnymO	9024	303	0	499	126542	94
2729969164	artemmobil	8857	394	3	678	148994	123
2817931146	123Lenchen	8835	473	3	722	153820	129
522645008	RaimundRaimund	8061	317	0	453	129619	115

Wobei sich die Tweets dieser Gruppe stark überschneiden (siehe dazu [3]). Man könnte auch diese sieben User anstatt von AfD-16 für die Klassifikation verwenden. Bei den Retweets müsste man jedoch nicht @mrstone0856&Co selbst sondern jene User, die von dieser Gruppe ihrerseits retweetet wird, verwenden. Dieser Ansatz wurde in [3] auch verfolgt. Der Vorteil ist: @mrstone0856&Co versenden wesentlich mehr Tweets als die Mitglieder von AfD-16. Sie decken ein wesentlich breiteres Spektrum der rechtspopulistischen Szene ab. Z.B. sind sie auch Fans der Österreichischen Identitären und insbesondere vom Identitären Obmann Martin Sellner. Das ist aber auch wieder der Nachteil dieser Methode. Darüber hinaus kann ein User wie @mrstone0856 über Nacht von Twitter verschwinden. Er ist außerhalb der Twitter-Community ein Nobody. Es kann auch ein AfD-Politiker sein Amt aufgeben. Allerdings erfährt man dies durch die traditionellen Medien und seine Position wird durch eine andere Person ersetzt. Die Analyse wird durch die in diesem Paper verwendete Methode wiederholbar und orientiert sich an politisch relevanten Personen. Wenn man mehr am breiten Spektrum rechtspopulistischer Stimmungen interessiert ist, dann bietet sich @mrstone0856 als „Musterhund“ an.

4895469501	AfD_Frankenthal	142	0	0	0	2028	93
4897268177	Saalejournal	33	0	0	10	485	98

Mit Hilfe der Schindel wird auch die AfD-Frankenthal und das @Saalejournal richtig klassifiziert. Beim @Saalejournal kommen zusätzlich AfD spezifische Hashtags hinzu.

814151634037129216	KREUZBERGERIN2	10	0	10	0	50	-100
--------------------	----------------	----	---	----	---	----	------

Der VaFCI wurde in dieser Untersuchung nicht für die Erkennung von prononcierten Gegner optimiert. Dies ist nur ein Nebenprodukt. Den geringsten VaFCI von -100 hat User @KREUZBERGERIN2. Die 10 Tweets im Juni sind ausschließlich Replies. Man kann darüber diskutieren, ob dieser User der aktivste AfD-Gegner ist. Dass er/sie auf die AfD nicht gut zu sprechen ist, geht aus den Tweets jedoch eindeutig hervor (siehe Grafik folgende Seite).

@Alice_Weidel Was ist mit dem Vorfall im Dom ? Auch ein Held oder gilt das nur wenn Polizisten auf Flüchtlinge und nicht auf Österreicher schießen?

@Alice_Weidel Wer hier scheitert, ist die AfD- keine Lösungen-nur Hetze und Hämte. Sie blamieren sich nicht entweder im Bundestag oder rennen raus - wollen Sie so die neue Bundeskanzlerin werden ? 🙄🙄🙄

@Alice_Weidel Ich kaufe ein nicht

@Beatrix_vStorch Vielleicht sollten Sie an ihrn Englisch arbeiten.. da steht keine Forderung, sondern eine Botschaft 🙄

@Beatrix_vStorch Gehen Sie doch einfach in ein anderes Land, wenn es Ihnen hier nicht gefällt - da es ja für Sie keine Klimaveränderung gibt, verstehen Sie diese Aktion natürlich auch mal wieder nicht ..

@Beatrix_vStorch Ach Frau von Storch - Sie langweilen mich - haben Sie auch was zu sagen ?

@Alice_Weidel Woran machen Sie den Erfolg fest - dass eine Partei, Minderheiten diskriminiert und ein Land mit Hetze gegen alles nicht Konforme nervt !?!

@Beatrix_vStorch Wer hier lügt hat ihr englischsprachiges Interview gezeigt - selten so gelacht -

@Beatrix_vStorch Ihre Vorschläge sind so schlecht, dass es bald jeder begriffen haben wird

@Beatrix_vStorch Die DDR anzuführen ist so ziemlich das dümmste, was Sie tun konnten. Halten Sie uns Deutsche für blöd ???

Der VaFCI ist nicht perfekt. So hat der User *@AfDLDS*, die AfD Ortsgruppe Dahme-Spreewald, einen Index von 0. Sie wird überhaupt nicht als solche erkannt. Die Ortsgruppe verwendet keine Retweets und Hashtags. Sie weist meist auf externe Links und insbesondere Videos hin. Vanellus ist in der Lage, diesen Links zu folgen. Dies ist auch ein nettes Feature des Programms. Man klickt einen Tweet an und es erscheint in einem Browser-Fenster der im Tweet enthaltene Link. Allerdings ist es hoffnungslos, auf diese Art und Weise 29,2 Millionen Tweets zu durchsuchen. Man müsste diese Information vorab mit einem Web-Crawler a la Google herunter laden und entsprechend analysieren. Dies ist eine Aufgabe, die die Grenzen des Projektes sprengen.

Machbar wäre jedoch eine Analyse der Metadata. Dies sind alle Informationen, die man mit der internen Kennung von Twitter erhält. Man würde z.B. wissen, dass es sich um die Ortsgruppe Dahme-Spreewald handelt. Man kann argumentieren, dass dies auch aus dem Usernamen *@AfDLDS* hervor geht. Es gibt jedoch z.B. den User *@AfD_Hesse*, der keine Orts- oder Landesgruppe ist. *@AfD_Hesse* ist ein Tierschützer, der primär PETA retweeted und nur sekundär Sympathien für die AfD erkennen lässt.

@ContraAfD1 oder *@KeineAfD2021* sind natürlich Gegner. Es ist für ein Programm nicht trivial, dies zu erkennen. Es kann sich auf Twitter jeder so nennen wie er will. Der Name muss nur eindeutig sein. Nur bei einem verifizierten Account kann man davon ausgehen, dass die Informationen auch richtig sind. Alice Weidel hat z.B. einen verifizierten Account, *@AfDLDS* ist hingegen nicht verifiziert.

Zusammenfassung und Ausblick:

Diese Arbeit zeigt, dass man Follower mit relativ einfachen Mittel mit hoher Treffsicherheit klassifizieren kann. Es entsteht durch die vorgestellten Methoden kein Aufwand für die – manuelle – Erstellung der Kriterien. Diese werden dynamisch mit Hilfe der Daten erzeugt. Eine mit vertretbarem Aufwand realisierbare Verbesserungsmöglichkeit ist die Einbeziehung der Metainformation (siehe [4]). Eine wesentliche Verbesserung der generellen Analysemöglichkeiten wäre die Einbeziehung von anderen Social-Media, insbesondere von Facebook. Man müsste die Twitter und Facebook Daten verknüpfen. Im Moment konzentriert sich die Entwicklung von Vanellus auf Twitter, eine Erweiterung auf Facebook ist mittelfristig geplant.

Durch die Abfrage der Metainformation mit Hilfe der internen Kennung und durch die Verlinkung auf andere soziale Medien, insbesondere Facebook, sollte es möglich sein, den User mit hoher Wahrscheinlichkeit zu identifizieren. Dies war jedoch weder die Aufgabenstellung dieser Untersuchung noch gehört es zu den generellen Zielen des Autors. Vanellus ist ein Tool zur Analyse von gesellschaftlichen Stimmungen und kein Mittel zur Rasterfahndung.

Externe Daten:

Es wurden für die Analyse 232.330 Follower verwendet. Davon sind 17.630 aktive User die im Juni 2018 mindestens 10 Tweets gesendet habe. Die VaFCI-Werte für diese aktiven User können von <http://www.godotfinance.com/Vanellus/csv/VaFCIAfD.csv> als *.csv Datei herunter geladen werden. Der Zeichensatz ist „Westeuropäisch – ISO-8859-1“.

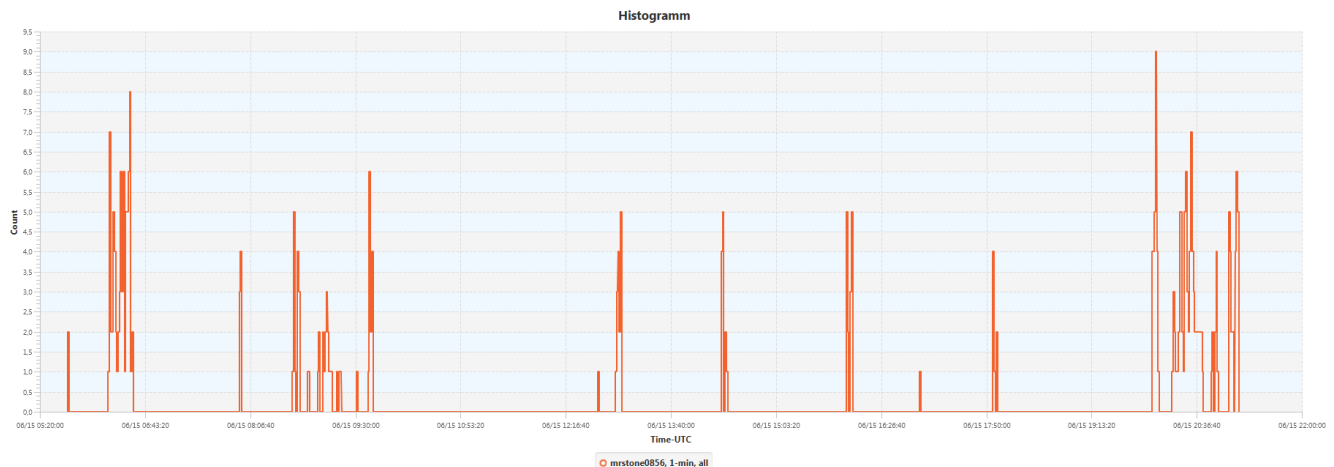
Von diesen 17.630 Usern haben 6.620 eine AfD-VaFCI von 60 und mehr. Diese User sind mit hoher Wahrscheinlichkeit AfD Anhänger. Die Daten dieser User können von http://www.godotfinance.com/Vanellus/csv/VaFCIAfD_min60.csv herunter geladen werden.

Verwendete Literatur:

- [1] A. Leavitt et. al.: The Influentials: New Approaches for Analyzing Influence on Twitter.
- [2] Ch. Donninger: Wie misst man den Einfluss von MdBs auf Twitter? Der Vanellus-Twitter-Index VaTI. <http://www.godotfinance.com/Vanellus/pdf/VaTIMdB.pdf>
- [3] Ch. Donninger: Der Twitterbot Vanellus zur vollständigen Analyse deutschsprachiger Tweets. <http://www.godotfinance.com/Vanellus/pdf/VanellusV029.pdf>
- [4] B. Perez, M. Musolesi, G. Stringhini: You Are Your Metadata: Identification and Obfuscation of Social Media Users using Metadata Information.
- [5] N. Chavoshi, H. Hamooni, A. Mueen: On-Demand Bot Detection and Archival System.

Anhang: Ist @mrstone0856 ein Bot oder ein Mensch?

@mrstone0856 und die übrigen User, die oben als „Rechtspopulistische Hardcore Community“ bezeichnet wurden, verfasst Monat für Monat 10.000 Tweets. Das ist – auf Dauer – für einen Menschen ungewöhnlich viel. Es könnte sich auch um einen Bot bzw. bei der gesamten Community um ein bot-network handeln. Es gibt verschiedene Methoden, bots zu erkennen. Laut [5] ist das einfachste und effektivste Kriterium das zeitliche Verhalten.



Die Grafik zeigt die Anzahl der Tweets/min von @mrstone0856 am Freitag, 15. Juni. Er hat frühmorgens ein intensives „Workout“ mit einer Spitzenleistung von 8 Tweets/min. Danach folgt eine gute Stunde Pause, mit der nächsten, nicht so intensiven Einheit. Das wiederholt sich mehrmals am Tag. Am Abend legt @mrstone0856 richtig los und erreicht dabei 9 Tweets/min. 9 Tweets/min ist für einen Menschen eine beachtliche Leistung. Es geht sich jedoch aus, wenn man nur retweeted (was @mrstone0856 meistens tut). Der Sprint wird nicht lange durchgehalten, die Tweetrategie sinkt nach kurzer Zeit wieder deutlich.

Es wurde auch die Korrelation des zeitlichen Verhaltens mit den übrigen Mitgliedern dieser Community berechnet. Man stellt eine hohe inhaltliche Überschneidung bei den Retweets fest, die zeitliche Koordinierung ist jedoch – für die Zeitdimension eines bot-networks – gering. Vom Zeitverhalten her besteht @mrstone0856&Co aus Menschen. Wenn @mrstone0856&Co ein bot-network ist, dann simuliert es nicht nur zeitlich sondern auch inhaltlich sehr gut menschliches Verhalten. Es ist plausibler, dass @mrstone0856&Co eine Gruppe von gleichgesinnten Menschen mit viel Tagesfreizeit sind.