

Universität Rostock

VERGLEICHENDE TWITTER-ANALYSE ÜBER DREI JAHRE HANSE SAIL – ORTE, EREIGNISSE UND STIMMUNGEN

Ferdinand Vettermann, Ralf Bill

Zusammenfassung: Im Rahmen des BMBF-Projekts KOGGE (Kommunale Gewässer Gemeinschaftlich Entwickeln) stehen neben hydrologisch-hydraulischen Fragestellungen auch die Verwendung innovativer Methoden zur Erfassung und Verarbeitung von Daten im Vordergrund. Einer dieser Ansätze ist die Integration von verortbaren Informationen aus sozialen Netzwerken wie Twitter. Um aufzuzeigen, welche Potenziale ein längerfristiges Monitoring für die Hanse- und Universitätsstadt Rostock bietet, wird in dieser Arbeit ein Vergleich zwischen den Großveranstaltungen der Hanse Sail 2016, 2018 und 2019 gezogen.

Pro Jahr konnten dabei etwa 1.000 Tweets in Rostock mit Bezug zur Hanse Sail verortet werden. Aus diesen wird deutlich, dass einzelne Ereignisse wie Unfälle in naher Echtzeit auf Twitter rezipiert und weiterverbreitet werden. Zum anderen zeigen sich deutliche Schwerpunkte, wobei diese sich über die Jahre zum Teil deutlich verändern.

Damit bietet Twitter für das Eventmanagement eine Möglichkeit, sowohl in naher Echtzeit Informationen zu erfassen und weiterzuverarbeiten, aber auch im Nachgang zu analysieren, wo besondere Interessen der Besucher lagen und wie die Stimmung in den diesbezüglichen Tweets war.

Schlüsselwörter: Soziale Medien, Twitter, Gazetteer-Matching, Lokationsidentifikation, Hanse Sail, Sentimentanalyse

COMPARATIVE TWITTER ANALYSIS OVER THREE YEARS HANSE SAIL – PLACES, EVENTS AND MOODS

Abstract: One of the core aspects of the BMBF project KOGGE was to solve hydrological and hydraulic problems with the help of innovative methods to gather and to process data. To achieve this goal, one idea is to harvest locatable information from social networks, especially Twitter. In order to show which potentials a longer-term monitoring offers for the Hanseatic and University City of Rostock, a comparison between the major events of the Hanse Sail 2016, 2018 and 2019 is drawn.

Per year more than 1,000 tweets could be located in Rostock with reference to the Hanse Sail. From these it becomes clear that individual events like accidents are received and spread near real time via Twitter. On the other hand, there are very clear annual focal points, although these are subject to some major differences.

Twitter thus offers for event management the opportunity to record and process information in near real time, but also to analyse afterwards where the special interests of visitors lay and what the mood in the tweets was like.

Keywords: Social media, Twitter, gazetteer-matching, location identification, sentiment analysis, Hanse Sail

Autoren

Dr.-Ing. Ferdinand Vettermann

Universität Rostock

Professur für Geodäsie und Geoinformatik

Justus-von-Liebig-Weg 6

D-18059 Rostock

E: ferdinand.vettermann@uni-rostock.de

Prof. Dr.-Ing. Ralf Bill

Universität Rostock

Professur für Geodäsie und Geoinformatik

Justus-von-Liebig-Weg 6

D-18059 Rostock

E: ralf.bill@uni-rostock.de

1 EINLEITUNG

Im Rahmen des BMBF-Projekts KOGGE (Kommunale Gewässer Gemeinschaftlich Entwickeln), in dem die Klein- und Kleinstgewässer sowie die Entwicklung eines integrierten Gewässerentwicklungskonzepts für die Hansestadt Rostock im Vordergrund stehen, befasst sich ein wesentlicher Teil mit der Nutzung innovativer Methoden der Datenerhebung und der Datenprozessierung sowie der Einbeziehung der Öffentlichkeit (Hübner & Vettermann 2016, Mehl et al. 2017, Tränckner et al. 2017).

Aus diesem Grund wurde geprüft, inwieweit soziale Medien, insbesondere Twitter, hier eine Möglichkeit bieten, Informationen zu erfassen und aufzubereiten, die sonst nicht verfügbar wären. Gerade Twitter mit seinen bis zu 500 Mio. Nachrichten pro Tag scheint sich hier außerordentlich zu empfehlen (Twitter Inc. 2015). So hat sich gezeigt, dass die Analyse von Tweets, vor allem in naher Echtzeit, eine Grundlage für die Entscheidungsunterstützung bieten kann (Vettermann & Bill 2019, Terpstra et al. 2012, Resch 2017). Neben globalen Fragestellungen, wie beispielsweise die Ausbreitung einer Grippewelle, lassen sich auch auf lokalem Maßstab wesentliche Informationen aus Twitter ableiten (Alkouz et al. 2019). Insbesondere in urbanen Räumen sind soziale Netze in der Lage, Informationen beispielsweise zur Mobilität und damit auch zur Verkehrsinfrastruktur zu liefern (Krüger et al. 2019).

Zudem bietet Twitter durch den einfachen Zugriff auf den Datenstrom eine ausgezeichnete Möglichkeit, freiwillig erhobene, raumbezogene und textuelle Informationen in großem Maßstab zu analysieren und zu verarbeiten. Dies unterstreichen zahlreiche Arbeiten (Hecht et al. 2011, Agarwal et al. 2011, Li et al. 2012). Allerdings gibt es ein deutliches Defizit an Forschungen diesbezüglich im deutschsprachigen Raum, insbesondere was den Aspekt der Entscheidungsunterstützung und Bürgerbeteiligung angeht (Fuchs et al. 2013, Scheffler 2014, Buschbaum et al. 2017).

Mit der ubiquitären Verfügbarkeit von Smartphones und der damit verbundenen Möglichkeit, Informationen (Texte, Bilder, Videos) von jedem Ort aus zu teilen und diese mit GNSS-Koordinaten zu versehen oder anders zu verorten, besitzen soziale Netzwerke einen Vorteil gegenüber den herkömmlichen Medien. Häufig sind so

Bildmaterial und Informationen über Ereignisse verfügbar, bevor die klassischen Medien überhaupt von deren Existenz wissen (Goodchild 2007). Durch diese auch als VGI (Volunteered Geographic Information) bezeichneten Daten lassen sich zudem Rückschlüsse auf das menschliche Verhalten ziehen, was insbesondere für die Citizen Science im Allgemeinen sowie die Umweltplanung im Speziellen wertvoll ist (Liu et al. 2015).

Die Potenziale, welche die großmaßstäbige Auswertung bei Veranstaltungen wie der Hanse Sail oder Demonstrationen bieten, wurden bereits an den Jahren 2016 und 2018 deutlich (Vettermann et al. 2017a, Vettermann & Bill 2019). Nun soll ein Vergleich zwischen der Sail 2019 zu den Vorjahren gezogen werden. Daraus sollen etwaige Änderungen bezüglich der Hotspots, der Themen, der Stimmungen sowie der Interessen der Besucher deutlich werden. Im Kern geht es darum, zu visualisieren, wie sich Hotspots und damit Interessen verschieben, was für die Einwohner von besonderer Wichtigkeit ist und damit schließlich zur Veranstaltungsplanung in den folgenden Jahren beizutragen.

2 ABLEITUNG VON ORTSBEZÜGEN AUS TWITTER BEI EREIGNISSEN

2.1 RAUMBEZOGENE TWITTER-ANWENDUNGEN

Soziale Netzwerke als Teil des Web 2.0 sind in jüngerer Zeit immer mehr in den Fokus von Forschung und Wissenschaft, aber auch von praktischen Anwendungen geraten. Es stehen hier vor allem die Citizen Science sowie die Nutzung der Daten als VGI im Vordergrund (Goodchild 2007, Haklay 2013). Dabei ist gerade die Ortsinformation, die sogenannte Location Based Social Networks (LBSNs) zu Nachrichten zur Verfügung stellen, von besonderer Bedeutung (Traynor & Curran 2013). Dadurch lässt sich der Inhalt einer Nachricht räumlich zuordnen und die Inhalte mit der Lokation verknüpfen. Die Methodik, die von einer großen Gemeinschaft freiwillig geschaffenen Datenbasis zu analysieren und zur Beantwortung verschiedenster Fragestellungen heranzuziehen, entspricht dem Grundprinzip des Crowdmings. Bei diesem werden die durch die Gemeinschaft erzeugten Daten ohne deren Wis-

sen, im Gegensatz zum Crowdsourcing, erhoben und ausgewertet (Streich 2014). Die frei verfügbaren Daten sozialer Netzwerke sind besonders für Fragen des Katastrophenmanagements, Analyse sozialer Milieus, der polizeilichen Einsatzplanung etc. wertvoll (Resch 2017, Longley & Adnan 2015, Terpstra et al. 2012). Die Anwendungen, die im räumlichen Bereich zum Einsatz kommen, stützen sich auf verschiedene Varianten. Neben der Ableitung der Ortsbezüge mithilfe eines Gazetteers oder mittels des sozialen Netzes und die Analyse der Beziehungen der Personen untereinander tritt immer mehr das maschinelle Lernen, insbesondere Deep Learning, in den Vordergrund (Huang et al. 2019, Vettermann et al. 2017b, Backstrom et al. 2010).

Neben den Methoden des maschinellen Lernens ist gerade bei Gazetteer-basierten Methoden der Vergleich zwischen Text und Wörterbuch notwendig. Bedingt durch die informelle Sprache bieten sich hier auf Fuzzylogik basierende Ansätze an (Vettermann et al. 2017b). Allerdings zeigen Recchia & Louwse (2013), dass Skip Grams in Kombination mit dem Dice-Koeffizienten zum Vergleich von deutschsprachigen Texten ideal erscheinen.

2.2 MASCHINELLES LERNEN IN DER COMPUTERLINGUISTIK

Wesentlicher Bestandteil bei der Verortung und Analyse von Tweets aber auch anderen Posts in sozialen Netzen ist die Verarbeitung des Texts. In der Regel kommen hier Ansätze des maschinellen Lernens zum Einsatz. Neben der Sentimentanalyse, auch Stimmungsanalyse genannt, ist vor allem die Identifikation von Trends als auch von Themen ein Bereich, wo diverse Verfahren des maschinellen Lernens zur Anwendung kommen (Cieliebak et al. 2017, Hendrickson et al. 2015, Benhardus 2013). Die Module umfassen dabei sowohl klassische Methoden des maschinellen Lernens wie Support Vector Machines (SVM) als auch Convolutional Neural Networks (CNN) oder Deep Learning. Im Rahmen der Textprozessierung kommen zudem häufig Recurrent Neural Networks (RNN), Long Short-term Neural Networks (LSTM) sowie Gated Recurrent Unit Networks zur Anwendung, welche z. T. äußerst performant sind (Bengio et al. 2006, Mikolov et al. 2011, Sundermeyer et al. 2013). Diese benöti-

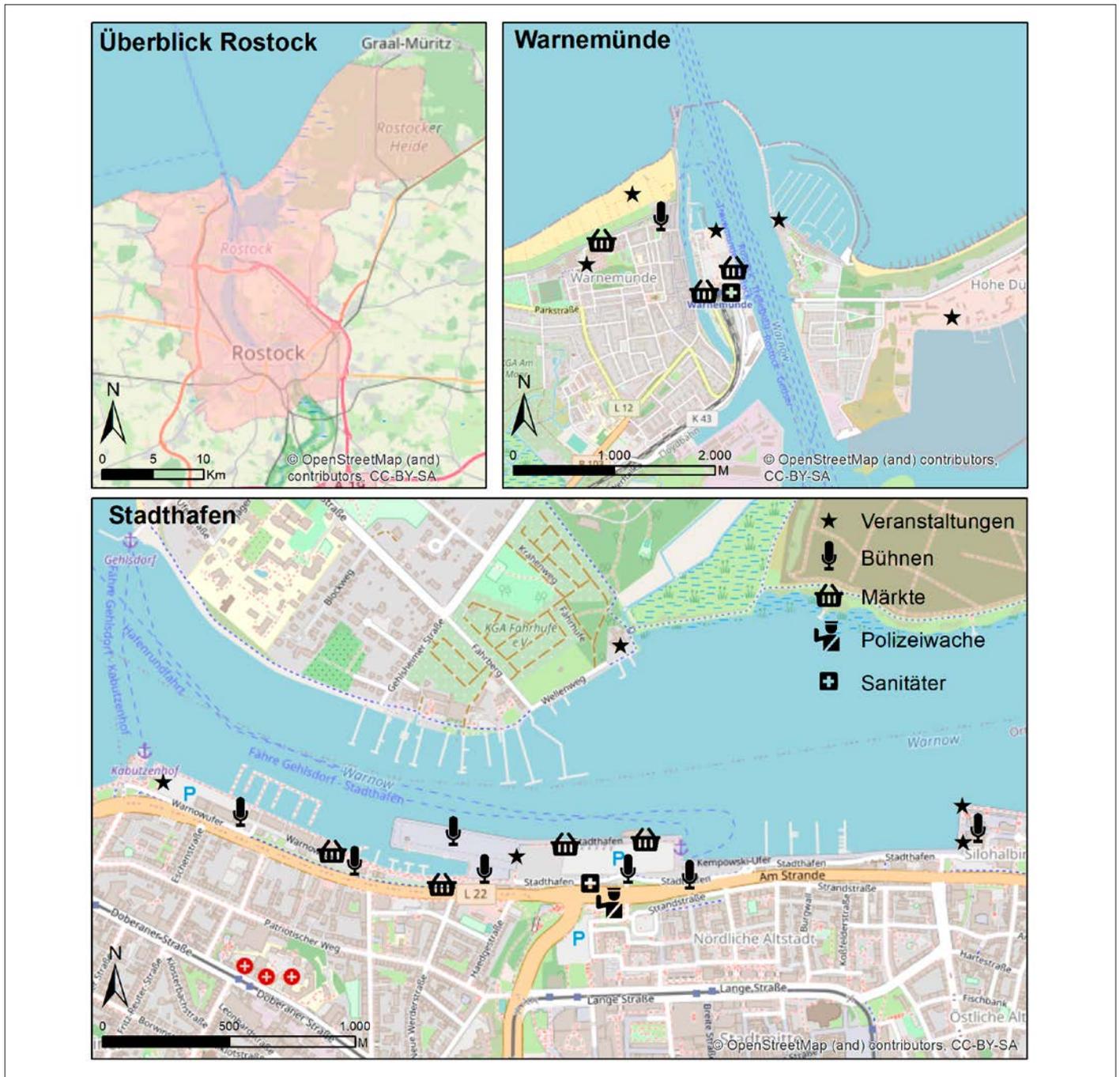


Abbildung 1: Veranstaltungsorte der Hanse Sail (Datengrundlage: OSM, OpenData.HRO, Hanse Sail 2019)

gen allerdings in der Regel viel Zeit zum Training, was den entscheidenden Nachteil dieser Netze darstellt. Demgegenüber lernen die aus der Bild- in die Textprozessierung transferierten CNNs nach Kim (2014) deutlich schneller.

Einer der bedeutendsten Anwendungsbereiche stellt die Identifikation von Trends und damit dominanten Themen in kurzen Zeiträumen dar. Twitters Algorithmus zur Trendidentifikation ist Teil deren Geschäftsgeheimnisses und stellt ein wesentliches Wettbewerbsselement dar. Daher ist dieser nicht öffentlich zugänglich. Das Erscheinen

in bestimmten Trends stellt beispielsweise für Marken eine gute Möglichkeit dar, explizit auf sich aufmerksam und damit Werbung zu machen (Twitter Inc. 2019, Bäckström & Haslum 2016). Allerdings erscheinen hier immer nur die Hashtags als Trends. Für die konkrete Anwendung ist dies hingegen weniger praktikabel, da auch abseits von Hashtags trendsetzende Themen identifiziert werden sollen. Grund dafür ist, dass viele Tweets lediglich unter dem Hashtag „Rostock“ geteilt werden, woraus aber letztlich kein konkretes Thema ermittelt werden kann.

Deshalb kommt in dieser Arbeit die Methodik der Latent Dirichlet Allocation (LDA) zur Anwendung. Das von Blei et al. (2003) entwickelte Verfahren ist eine der potentesten Methoden, Informationen aus Texten zu extrahieren (Wold et al. 2016). Die LDA behandelt hierbei jedes Dokument als einen Vektor bestehend aus der Anzahl der Wörter (BoW). Jedes Dokument beinhaltet dabei diverse Themen, wobei das einzelne Thema wiederum eine Wahrscheinlichkeitsverteilung über die Anzahl von Wörtern ist. Der Algorithmus der LDA weist jedem Dokument, in diesem Fall also jedem

Tweet, ein Thema entsprechend der multinominalen Verteilung der enthaltenen Wörter zu (Blei et al. 2003).

Des Weiteren gibt es auch Methoden, um Ortsbezüge aus Texten maschinell zu vergleichen. Dies ermöglicht es, vom statischen Gazetteer abzuweichen und die Verortung deutlich einfacher übertragbar zu machen (Hahmann et al. 2014). Inzwischen treten im Bereich der räumlichen Analyse verstärkt Verfahren des Deep Learnings in den Vordergrund. Eine der aktuell wohl performantesten Open-Source-Lösungen in diesem Bereich stellt wohl DeepGeo dar (Snyder et al. 2019, Lau et al. 2017).

2.3 DIE HANSE SAIL IN ROSTOCK

Die Hanse Sail als größtes Event in Mecklenburg-Vorpommern zählt zu den größten Volksfesten in Deutschland und weltweit zu den bedeutendsten Segelveranstaltungen. Das jedes Jahr über das zweite Augustwochenende stattfindende fünftägige Event (Donnerstag bis Sonntag, wobei am Mittwoch bereits das Einlaufen der Schiffe stattfindet) lockt jährlich etwa eine Million Besucher in die Hanse- und Universitätsstadt Rostock (Tourismuszentrale Rostock & Warnemünde 2018). 2019 fand die 29. Hanse Sail seit ihrer Gründung unter dem Namen Hanseatische Hafentage (Süddeutsche Zeitung 2019) durch den ehemaligen Oberbürgermeister Roland Methling im Jahr 1991 statt.

Rostock ist durch seine Lage an der Ostsee sowie der Warnowmündung prädestiniert, eine derartige Veranstaltung auszurichten, da es möglich ist, sie sowohl innenstadtnah als auch küstennah durchzuführen. Die Veranstaltung ist dabei vor allem im Stadthafen und in Warnemünde lokalisiert (Abbildung 1). Mit mehreren Bühnen, zahlreichen Ständen und diversen Fahrgeschäften stellt der Stadthafen die Hauptanlaufstelle für die Besucher dar. Zum anderen lockt auch Warnemünde die Zuschauer nach Rostock, wo unter anderem die Großsegler anlegen, da für deren Tiefgang und Größe der Stadthafen nicht ausreichend ist. Des Weiteren finden weitere Veranstaltungen entlang der Warnow statt (z. B. das Kanonier- und Böllertreffen in Gehlsdorf) (Tourismuszentrale Rostock & Warnemünde 2018).

Für die Reichweite ist für die Organisation der Hanse Sail von großer Bedeutung,

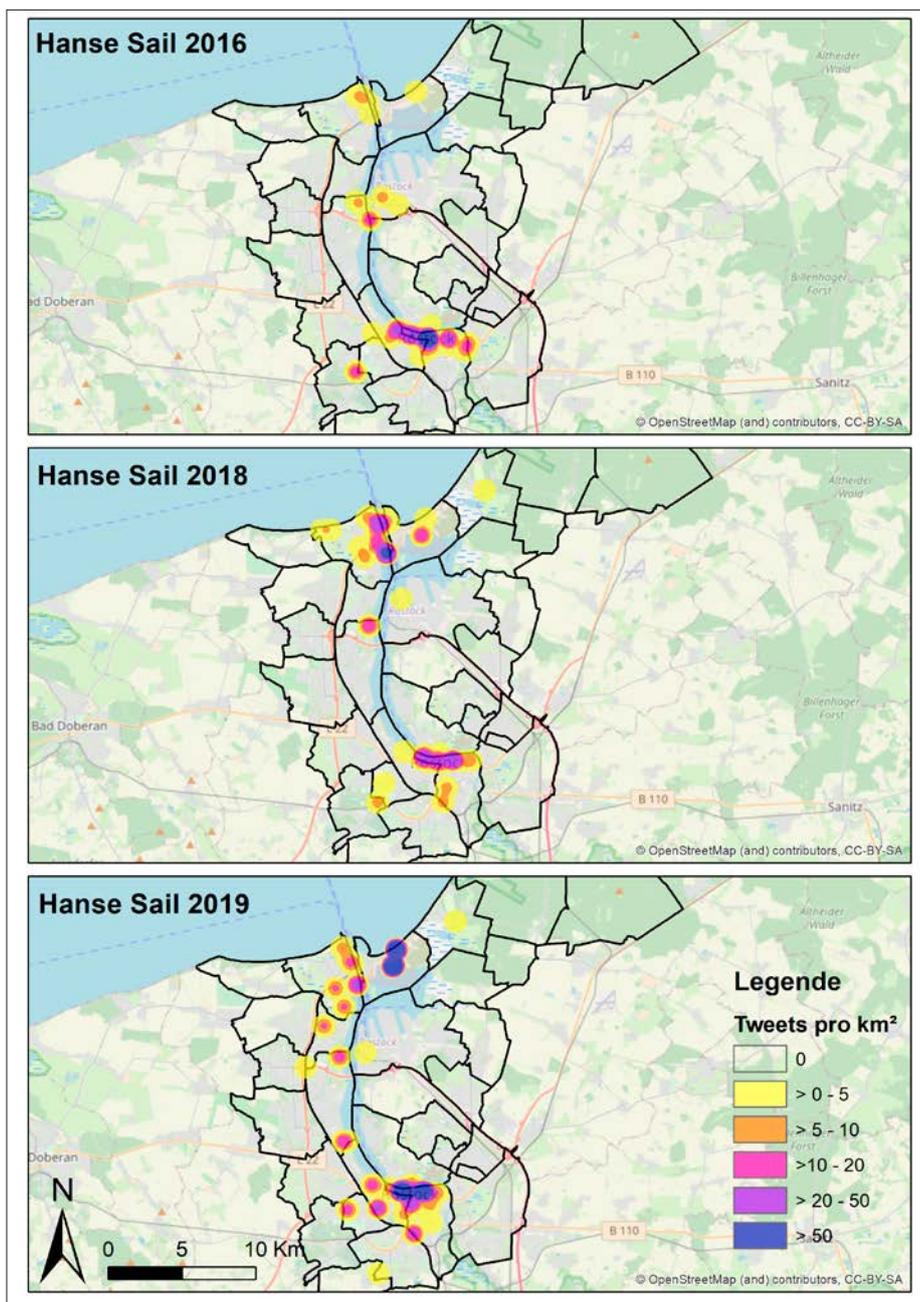


Abbildung 2: Kerndichte-Schätzung der Tweets, welche sich mindestens auf Straßenebene verorten ließen (Datengrundlage: OSM, OpenData.HRO, Twitter Inc.)

zu ermitteln, welches Interesse die Besucher an bestimmten Schiffen oder einzelnen Konzerten respektive Bühnen haben. Daneben ist es auch für das Sicherheitskonzept relevant, wo mehr Einsatzkräfte vorgehalten werden müssen, wo mit vermehrtem Zuschaueraufkommen zu rechnen ist und welche Events besonders beliebt sind.

3 DATENERFASSUNG

Die Erfassung der Tweets basiert auf der kostenfreien Twitter-Streaming-API. Durch diese ist es möglich, 1% des globalen Datenstroms abzugreifen. Für eine nur auf Rostock bezogene Anwendung reicht dies

allerdings, trotz der Großveranstaltung Hanse Sail, aufgrund des vergleichsweise geringen Tweetaufkommens in diesem Raum aus. So haben Vettermann & Bill (2019) gezeigt, dass in einem Zeitraum von sechs Wochen nur ein Maximum von 1681 Tweets pro Tag erreicht worden ist.

Der Zugriff auf den Stream selbst fand über die Python-Bibliothek Twython statt, wobei die Wahl auf Python aufgrund der Vielzahl an verfügbaren Modulen zum Machine Learning als auch zur sonstigen weiteren Datenverarbeitung fiel (McGrath 2018). Hierbei ist vor allem die Bibliothek NLTK (Sprachprozessierung), ScikitLearn

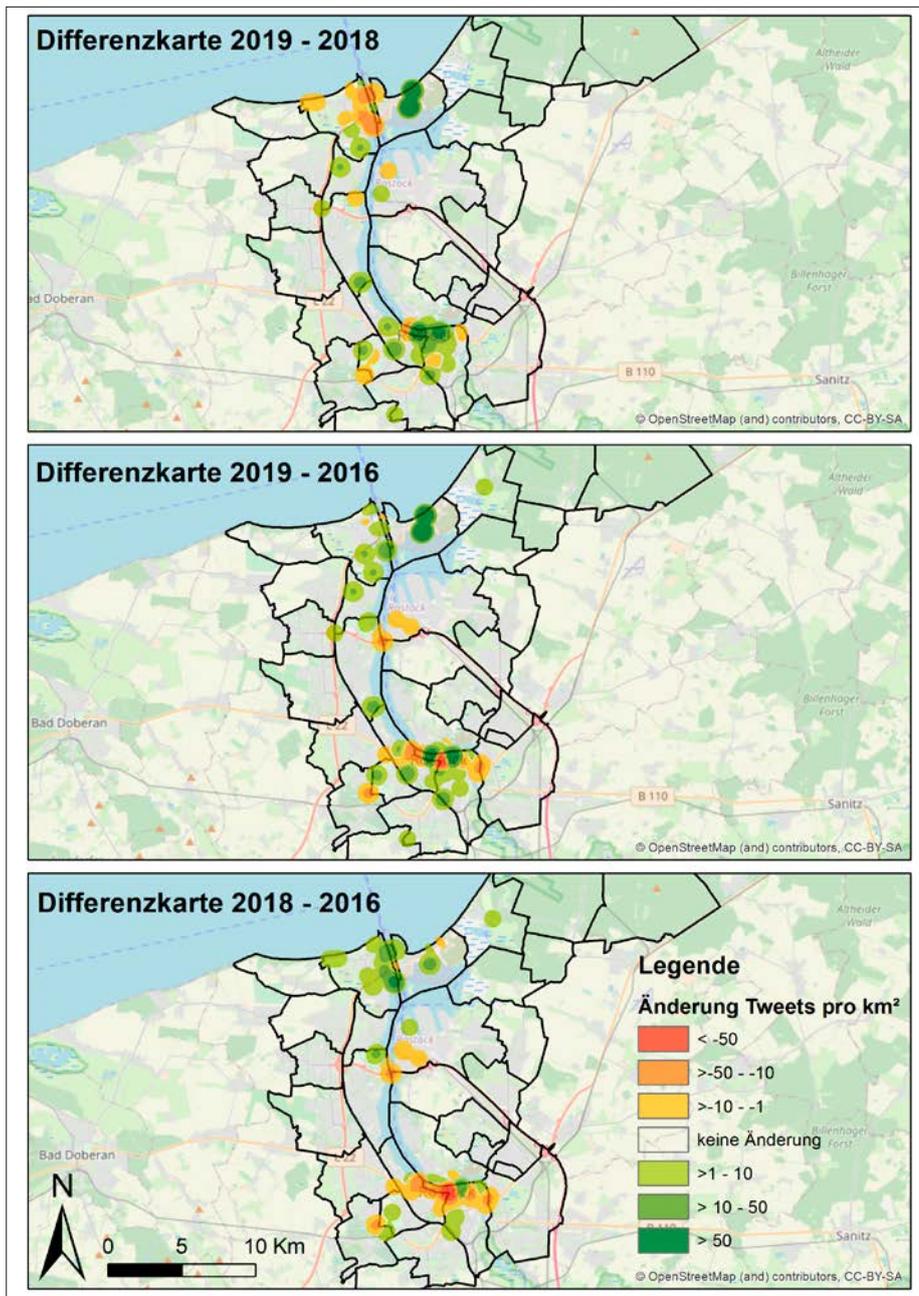


Abbildung 3: Differenzkarte der Tweetdichten für die drei Untersuchungszeiträume (Datengrundlage: OSM, OpenData.HRO, Twitter Inc.)

(Algorithmen zum maschinellen Lernen) sowie psychog2 (Datenbankverbindung zu PostgreSQL) hervorzuheben (Bird et al. 2009, Buitinck et al. 2013, Gregorio & Varrazzo 2018). Die Daten werden im leicht zu verarbeitenden JSON-Format zurückgeliefert. Jedem Tweet werden damit zugehörige Metainformationen wie Zeitzone, Follower oder die Anzahl der Retweets angehängt. Wesentlich für die Verortung ist der mehrschichtige Gazetteer. Mit dessen Hilfe werden die einzelnen Orte nach ihrer Genauigkeit verortet. Dabei sind die Einträge im Gazetteer von Ebene 0 (Stadt) über Ebene 1 (Stadtgebiete), Ebene 2 (Stra-

ßen) und Ebene 3 (Points of Interest) zugeordnet. Der Aufbau der gesamten Anwendung soll an dieser Stelle nicht weiter betrachtet werden, vielmehr sei hierzu auf Vettermann et al. (2018) verwiesen.

Neben der Erfassung spielen auch die Aufbereitung und Bereinigung der Texte eine wichtige Rolle. Dabei werden Buchstabenwiederholungen, Lachen oder Smilies entsprechend gefiltert und getaggt. Eine detaillierte Beschreibung aller Prozesse, mit denen die unkonventionelle und orthografisch meist inkorrekte Twitter-Sprache aufbereitet wird, findet in Vettermann (2019) statt (Miller & Goodchild 2015).

Im Rahmen dieser Arbeit werden den Tweets Stimmungen, kategorisiert in Positiv oder Negativ, mittels einer SVM-Klassifikation zugewiesen. Zum Training des Klassifikators wurden 150.000 positive und 150.000 negative deutschsprachige Tweets mittels positiver und negativer Smilies gesammelt. Der Klassifikator erreichte dabei eine Gesamtgenauigkeit von 83% (Vettermann 2019).

Zusätzlich zum Sentiment ist für die erfassten Nachrichten eine LDA durchgeführt worden, für die drei Klassen gewählt wurden. Grund für die relativ geringe Klassenanzahl ist die schlechte Trennbarkeit der Themen aufgrund der Ähnlichkeit (Verwendung der gleichen Schlagwörter wie Rostock, Hanse Sail, Schiffe, Wasser etc.) als auch die insgesamt relativ geringe Anzahl an Tweets.

4 ERGEBNISSE

4.1 RÄUMLICHE VERTEILUNG

Für jedes der drei betrachteten Jahre wurden die Tweets mit der Applikation gefiltert und anschließend entsprechend nach den Schlagwörtern „sail“ und „schiff“ der Thematik Hanse Sail zugeordnet. Als Zeitraum wurden jeweils fünf Tage (Mittwoch bis einschließlich Sonntag) betrachtet. Damit wird es möglich, eine vergleichbare Datenbasis zu schaffen, die anschließend nach einer Kerndichte-Schätzung mit ArcMap räumlich visualisiert worden ist (Abbildung 2). Eingeflossen sind dabei jeweils die Tweets, welche mindestens auf Straßenebene verortet werden konnten (Ebene 2 und 3).

Für die Hanse Sail 2016 fand dies bereits in Vettermann et al. (2017a) statt. Allerdings sind für eine bessere Vergleichbarkeit die Methoden zur Datenaufbereitung als auch die Zeiträume angepasst und weiterentwickelt worden. Daher weichen die vorgestellten Daten von den in Vettermann et al. (2017a) publizierten Ergebnissen leicht ab. Zudem wird in der vorliegenden Arbeit nur Bezug auf die Hanse Sail genommen, jedoch nicht auf weitere Ereignisse und Themen, die in diesem Zeitraum stattfanden.

Insgesamt sind 2016 im Zeitraum vom 10.08. bis zum 15.08.2016 919 Tweets in Rostock mit Bezug zur Hanse Sail verortet worden. Aus diesen ließen sich 107 Tweets auf den Genauigkeitsebenen 2 und 3 verorten und flossen damit in die in Ab-

Abbildung 2 dargestellte Kerndichte-Schätzung mit ArcMap ein, um die Tweets pro km² darzustellen (Suchradius: 500 m, Auflösung 10 m). Aus der Verteilung wird deutlich, dass vornehmlich über die Hanse Sail im Raum des Stadthafens getwittert wurde. Daneben sind kleinere Hotspots im Bereich des Überseehafens als auch in Warnemünde zu finden.

2018 sind im Zeitraum vom 08.08. bis zum 12.08.2018 insgesamt 819 Tweets aufgelaufen, von denen sich 102 auf den beiden genauesten Ebenen verorten ließen. Aus der abgeleiteten Kerndichte-Schätzung wird deutlich, dass die Schwerpunkte vor allem im Bereich des Stadthafens als auch in Warnemünde zu finden sind.

2019 sind insgesamt 1.170 Tweets im Zeitraum vom 09.08. bis 11.08.2019 aufgelaufen, von denen sich 281 auf den beiden genauesten Ebenen verorten ließen. Aus der Verteilung wird deutlich, dass die Tweets sich 2019 vor allem in Bereich des Stadthafens als auch beim Marinestützpunkt Hohe Düne konzentrieren. Zudem sind sie deutlich breiter im Innenstadtbereich verteilt. Im Raum Warnemünde standen vor allem die beiden Segelschiffe „Kruzenshtern“ und „Mir“ im Fokus der Aufmerksamkeit.

Um einen direkten, räumlichen Vergleich zwischen den drei Jahren zu ermöglichen, sind Differenzkarten aus den Kerndichte-Schätzungen berechnet worden (Ab-

Abbildung 3). Zwischen 2018 und 2016 wird deutlich, dass vor allem im Bereich Warnemünde mehr Nachrichten, im Innenstadtbereich hingegen weniger Nachrichten verortet werden konnten. 2019 gegenüber 2016 zeigt sich ein ähnliches Bild, obgleich im Bereich des Stadthafens auch Bereiche existieren, denen deutlich mehr Aufmerksamkeit auf Twitter entgegengebracht wird. Daneben fällt vor allem der Marinestützpunkt Hohe Düne auf, über den viel mehr getwittert wird als noch 2016. Beim Vergleich zwischen 2019 und 2018 bestätigt sich dieses Bild. Zudem wird sichtbar, dass 2018 bei den drei Vergleichsjahren vor allem Warnemünde im Fokus lag. Alles in allem kann man also von einer räumlichen Verschiebung der Interessen von 2016 (Stadthafen) über 2018 (Warnemünde) zu 2019 (Hohe Düne) sprechen.

4.2 ZEITLICHER VERLAUF

Daneben ist für jedes Jahr der zeitliche Verlauf der Nachrichten über die fünf Tage (Abbildung 4) dargestellt. In diesen Verlauf sind alle Tweets mit Bezug zur Sail eingeflossen, nicht nur die auf der Heatmap verorteten. Aus dem sich daraus ergebenden Graphen lassen sich z. T. sehr deutlich einzelne Ereignisse der jeweiligen Veranstaltungen erkennen. Bei der Hanse Sail 2016 fällt insbesondere ein Peak in der Nacht zwischen Samstag und Sonntag auf. Die-

ser ist auf eine große Anzahl an Tweets über die Japanischen Zerstörer Asagiri und Setuyoki zurückzuführen, welche bereits am Freitag der 2016er Hanse Sail im Überseehafen angelegt haben. Interessant ist, dass hierbei eine große Zahl japanischer Accounts eine Nachricht von „Rostock Heute“ aufgegriffen und offenbar einen Tag später in großer Anzahl (insgesamt 133 Tweets) getweetet haben.

2018 fallen vor allem zwei Peaks ins Auge. Der erste befindet sich am Abend des zweiten Tags, der zweite abends und vormittags des dritten respektive vierten Tags der Sail. Der erste Peak ist durch die Unterbrechung der Hanse Sail aufgrund einer Sturmwarnung entstanden (Nordkurier 2018). Der diesbezügliche Tweet wurde von vielen Accounts aufgegriffen und getweetet und konnte damit eine entsprechende Reichweite entfalten. Der zweite Peak lässt sich auf die Beschädigung der „Mir“ durch ein Sportboot am Freitagabend zurückführen (Ostseezeitung 2018). Dieses Ereignis stieß auf entsprechendes Interesse und fand daher auch auf Twitter eine entsprechende Resonanz.

2019 schließlich sind keine besonderen Peaks zu erkennen. Es wird allerdings deutlich, dass sowohl der Donnerstag als auch der Freitag der 2019er Sail auf Twitter deutlich mehr Beachtung fanden. Allerdings sind die Nachrichten nicht auf Einzelereignisse zurückzuführen. Auffällig ist aller-

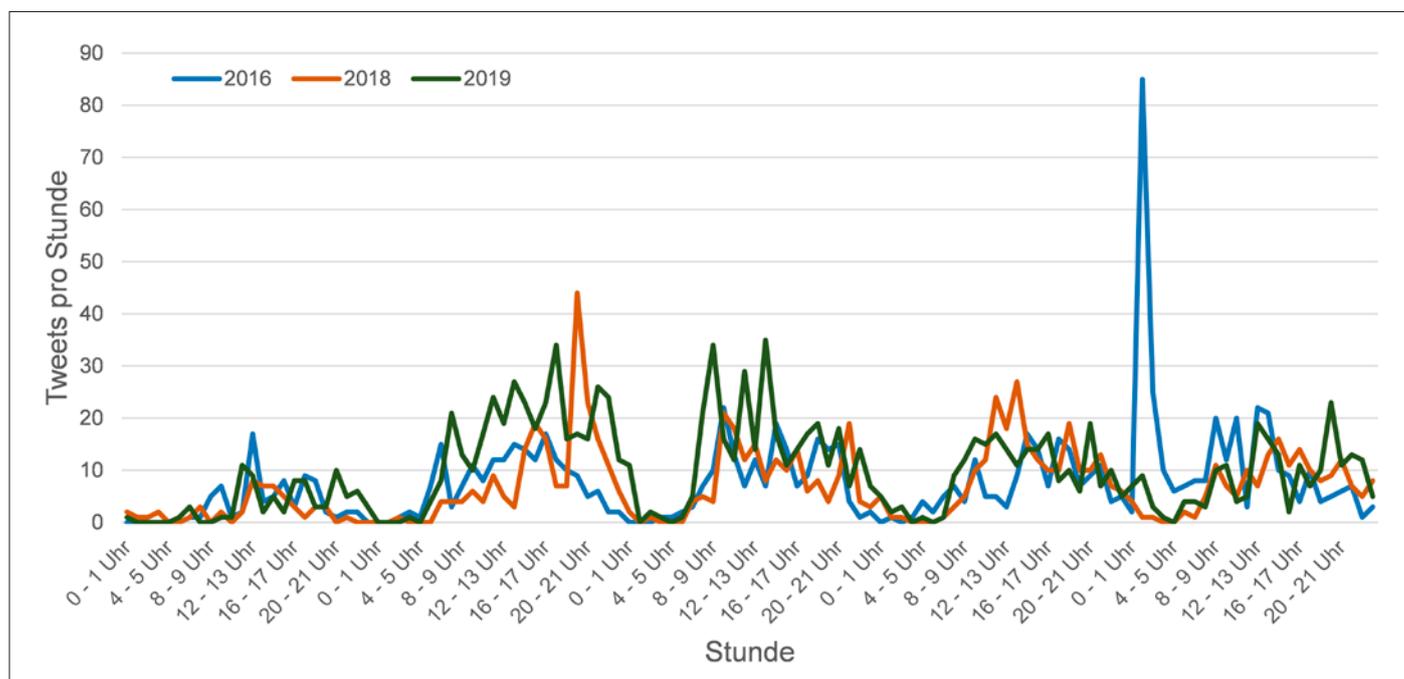


Abbildung 4: Zeitlicher Verlauf der Tweets in den drei Untersuchungszeiträumen

Thematik	Schlagworte 2016	Schlagworte 2018	Schlagworte 2019
1	Polizei, Warnemünde, Bühne Antenne MV	Warnemünde, Seebrücke, Achtung, Sturm	Hohe Düne, Gloria, Deutsche Marine, Willkommen
2	Segelstadion, Schiff, Ahoi	Freitag, Besucher, Feuerwerk, Wetterlage	Warnemünde, Wochenende, Flaggschiff, Fregatte
3	Japanische Zerstörer, Asagiri, Open Ship	Videos, Ostsee, Party, Ganztägig	Stadthafen, mexikanische, Besu- cher, Cuauhtemoc, Seenotrettung
Kohärenz	0,44	0,49	0,47

Tabelle 1: Schlagworte der Themen abgeleitet nach LDA im Untersuchungszeitraum

dings, dass eine Vielzahl an Tweets Bezug zum Marinesützpunkt Hohe Düne aufweisen. Dieser stand offenbar bei der Sail 2019 im Vordergrund des Interesses und wurde auch entsprechend durch institutionelle Accounts beworben.

4.3 THEMATIKEN

Um die Tweets der drei Hanse Sails entsprechend kategorisieren zu können, wurde eine LDA mit drei Klassen ausgeführt. Eine Prüfung auf eine optimale Anzahl an Klassen ergab, dass eine gute Kohärenz erst ab fünf Klassen oder aber bei einer Klasse gegeben ist. Aus diesem Grund wurde sich für einen Mittelweg entschieden, um sowohl der relativ kleinen Datenbasis aber auch der erwarteten Zergliederung in bestimmte Themen gerecht zu werden. Die mittels LDA abgeleiteten, dominanten Themen sind in Tabelle 1 dargestellt. Die abgeleiteten Themen wurden hierbei um die Schlagworte Rostock, Hanse Sail und Schiff bereinigt.

2016 lassen sich die Themen grob in allgemeine Veranstaltungen, konkrete Segel-events sowie die Besichtigungsmöglichkeit der japanischen Zerstörer einteilen. Allerdings ist die Trennbarkeit der Klassen relativ schlecht, mit Ausnahme der japanischen Schiffe. Dies wird auch durch die Kohärenz von 0,44 unterstrichen. 2018 beschreibt die erste Thematik Aktionen des Vereins „Seebrücke“ als auch die Sturmwarnungen. Dies überlagert sich allerdings stark mit der zweiten Kategorie, welche ebenfalls Bezug zur Wetterlage aufweist. Lediglich die dritte Thematik ist hinsichtlich ihrer Schlagworte recht gut abgrenzbar. Sie bezieht sich vor allem auf die Sail als Veranstaltung insgesamt. 2019 schließlich weist die besten Trennbarkeiten zwischen den drei Themen

auf. Thema eins bezieht sich vor allem auf Hohe Düne und beinhaltet auch das Schiff Gloria. Thema zwei bezieht sich auf Warnemünde und die dort auffindbaren Fregatten. Thema drei schließlich umfasst die Schiffe und Veranstaltungen im Stadthafen.

4.4 SENTIMENTANALYSE

In die SVM basierte Sentimentanalyse, kategorisiert in Positiv und Negativ sowie deren jeweilige Zuordnungswahrscheinlichkeit, sind alle korrigierten Tweettexte der jeweiligen Jahre mit Bezug zur Sail eingeflossen (vgl. Kapitel 3). Für die Hanse Sail 2016 ergibt sich unter Berücksichtigung von einer absoluten positiv/negativ Zuordnung ein mittleres Sentiment von 0.81, also stark positiv. Die mittlere positive Zuordnungswahrscheinlichkeit liegt immer noch bei 66%. Folglich wird über die Sail 2016 sehr stark positiv getwittert.

2018 liegt der mittlere Wert bei der absoluten Zuordnung bei 0.79, die Zuordnungswahrscheinlichkeit für positiv bei 67%. Damit fallen die Tweets zur Sail 2018 immer noch stark positiv aus, der Anteil ist dabei vergleichbar.

2019 schließlich liegt die absolute Zuordnung bei 0.72, die Zuordnungswahrscheinlichkeit für positiv bei 66%. Damit wird weniger positiv über die Sail 2019 berichtet als in den Vorjahren. Nichtsdestotrotz ist die Berichterstattung erneut stark positiv.

5 DISKUSSION

Um die Ergebnisse einordnen sowie beurteilen zu können, ist eine kritische Betrachtung notwendig. Kritisch anzumerken ist, dass das Verfahren der Erfassung für die Hanse Sail 2016 noch nicht die Qualität hatte, wie die der Jahre 2018 und 2019.

Ursächlich dafür sind zahlreiche Erweiterungen und Verbesserungen, die aufgrund der ersten Ergebnisse im Jahr 2016 in die Software eingeflossen sind (Vettermann et al. 2017a, Vettermann & Bill 2019).

Insbesondere die deutliche Erweiterung des Gazetteers ist dabei ausschlaggebend, aber auch die Verbesserung des Matching-Algorithmus selbst, der nicht mehr auf reiner Fuzzy-Logik, sondern auf Skip-Grams in Kombination mit dem Dice-Koeffizienten basiert. Hier hat sich in anderen Arbeiten gezeigt, dass dieses Verfahren vor allem im Deutschen ideal für den Textvergleich erscheint (Recchia & Louwerse 2013).

Eine Verschiebung der Hotspots über die fünf Tage des jeweiligen Untersuchungszeitraums (jeweils Mittwoch 0 Uhr bis Sonntagabend 23:59 Uhr) ist ebenfalls zu erkennen. Grundsätzlich bleiben allerdings der Stadthafen und Warnemünde die Hauptanlaufstellen für die Besucher. Eine gewisse Ausnahme stellt der Marinesützpunkt Hohe Düne 2019 dar. Ursächlich dafür scheint vor allem die Werbung der Bundeswehr für die Besichtigung der Schiffe im Marinesützpunkt zu sein, die 2019 stärker ausfiel als in den Vorjahren.

Im Jahr 2018 zeigt sich, dass das Werkzeug auch in der Lage ist, Einzereignisse in naher Echtzeit zu erfassen. So kam es hier erneut zu einem Unfall (2017 ist bereits der Eisbrecher „Stettin“ mit einer Fähre zusammengestoßen), bei dem das Segelschulschiff „Mir“ mit einem Motorboot kollidierte (Ostseezeitung 2018). Damit zeigt sich, dass Twitter nicht nur bei Großereignissen wie Erdbeben wertvolle Informationen in naher Echtzeit liefern kann, sondern auch bei Großveranstaltungen wie der Hanse Sail in der Lage ist, In-

formationen bereitzustellen. Dies untermauert die in Vettermann & Bill (2019) aufgestellte These mit Bezug zur Demonstration der AfD in Rostock im September 2018.

Die Zuordnung von Themen mittels LDA zeigt nur durchschnittliche Ergebnisse (Röder et al. 2015). Grund dürfte hier die relativ geringe Nachrichtenzahl sein, die die Differenzierung der einzelnen Themen erschwert. Zudem spielt hier der Faktor eine Rolle, dass im Grunde nach Subthemen gesucht wird, da das Hauptthema mit der Hanse Sail bereits festgelegt worden ist. Eine LDA aller Tweets bezüglich Rostock im Untersuchungszeitraum könnte hier eventuell bessere Ergebnisse bringen, wobei eine Ableitung von Subthemen hier dieselben Probleme aufweisen würde.

Interessant ist, dass über die Hanse Sail extrem positiv berichtet wird. Vergleicht man dies mit den Ergebnissen von Vettermann (2019) wird deutlich, dass über die Hanse Sail selbst deutlich positiver berichtet wird als über Rostock insgesamt. Allerdings ist anzumerken, dass über die 2019er Sail weniger positiv geschrieben wird als dies in den Vorjahren der Fall war. Daneben wäre eine Analyse mit weiteren Verfahren, wie CNNs, eine interessante Möglichkeit, die Ergebnisse zu verifizieren (Cieliebak et al. 2017).

6 ZUSAMMENFASSUNG UND AUSBLICK

Twitter bietet eine Möglichkeit, raumzeitliche und thematische Informationen in hoher Qualität und naher Echtzeit zu erheben und bereitzustellen. Dies gilt auch für kleinere Städte wie Rostock, wobei gerade bei Großveranstaltungen Twitter seine Stärke ausspielen kann. Dies wird besonders deutlich an der Hanse Sail in der Hanse- und Universitätsstadt Rostock mit ihren jährlich etwa 1 Mio. Besuchern.

Durch das hier vorgestellte Verfahren wird es möglich, sowohl a posteriori als auch in naher Echtzeit Ereignisse, Interessen und Hotspots der Hanse Sail zu identifizieren. Des Weiteren ist es möglich, Veränderungen der Interessenlage der Twitter-Community bezüglich der Sail zu visualisieren. Daneben ergibt sich aus der Themenanalyse ein Bild, über was die Twitter-Community während der Sail spricht. Vor allem das sehr positive Sentiment zeigt aber, wie beliebt die Hanse Sail ist, obgleich es einer längeren Zeitreihe bedarf,

die Stimmungsveränderungen genauer zu analysieren. Zudem wäre eine Analyse mit Verfahren wie CNNs eine wertvolle Ergänzung und würde eine bessere Vergleichbarkeit der Ergebnisse ermöglichen.

Wesentlich für die Analyse ist zudem die erfolgreiche Lokalisation der Nachrichten. Diese basiert vor allem auf den Einträgen im Gazetteer, folglich ist sie entsprechend von deren Anzahl und den entsprechenden Begriffen abhängig. Damit wird allerdings eine Übertragbarkeit erschwert. Aus diesem Grund wäre es eine gute Ergänzung, den Gazetteer auf Basis eines Webservices zu erstellen, der neue Ortsbezeichnungen automatisch integriert.

Folglich reichen Tweets alleine in der Regel nicht aus, komplexe Entscheidungen im Rahmen des Veranstaltungsmanagements zu treffen. Allerdings stellen sie eine hervorragende, zusätzliche Ressource zur Informationsgewinnung dar. Vor allem Unsicherheiten in deren Analyse machen weitere Informationen unersetzlich für die Entscheidungsfindung.

DANKSAGUNG

Die Verfasser danken dem Bundesministerium für Bildung und Forschung für die Förderung im Rahmen des ReWaM-Projekts KOGGE (FKZ: 033W032A).

Literatur

- Agarwal, A.; Xie, B.; Vovsha, I.; Rambow, O.; Passonneau, R. (2011): Sentiment Analysis of Twitter Data. In: Nagarajan, M.; Gamon, M. (Eds.): Proceedings of the Workshop on Languages in Social Media. Association for Computational Linguistics, Stroudsburg, S. 30-38.
- Alkouz, B.; Aghbari, Z. A.; Abawajy, J. H. (2019): Tweetluenza: Predicting flu trends from twitter data. In: Big Data Mining and Analytics, 2 (4), S. 273-287.
- Backstrom, L.; Sun, E.; Marlow, C. (2010): Find me if you can. In: Rappa, M. (Ed.): Proceedings of the 19th International Conference on World Wide Web: WWW'10, 26. – 30.04.2010, Raleigh, NC. ACM Digital Library, 61, New York.
- Bäckström, S.; Haslum, J. F. (2016): Detecting Trends on Twitter: The Effect of Unsupervised Pre-Training. Stockholm.
- Bengio, Y.; Schwenk, H.; Senécal, J.-S.; Morin, F.; Gauvain, J.-L. (2006): Neural Probabilistic Language Models. In: Holmes, D. E.; Jain, L. C. (Eds.): Innovations in Machine Learning: Theory and Applications. Studies in Fuzziness and Soft Computing, 194. Springer, Berlin/Heidelberg, S. 137-186.
- Benhardus, J. (2013): Streaming Trend Detection in Twitter. In: International Journal of Web Based Communities, 9 (1), S. 122-139.
- Bird, S.; Klein, E.; Loper, E. (2009): Natural Language Processing with Python. O'Reilly, Sebastopol.
- Blei, D. M.; Ng, A. Y.; Jordan, M. I. (2003): Latent Dirichlet Allocation. In: Journal of Machine Learning Research, 3, S. 993-1022.
- Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; Layton, R.; Vanderplas, J.; Joly, A.; Holt, B.; Varoquaux, G. (2013): API design for machine learning software: Experiences from the scikitlearn project. In: ECML PKDD Workshop: Languages for Data Mining and Machine Learning, S. 108-122.
- Buschbaum, K.; Blitz, A.; Reithmeier, C.; Kanwischer, D. (2017): Hashtags und Raumkonstruktionen: Eine explorative Studie zum Potential von digitalen Methoden zur Analyse raumzeitlicher Daten in sozialen Medien. In: gis.Science, 4/2017, S. 115-125.
- Cieliebak, M.; Deriu, J. M.; Egger, D.; Uzdilli, F. (2017): A Twitter Corpus and Benchmark Resources for German Sentiment Analysis. In: Lun, W. K.; Cheng, T. L. (Eds.): Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, Valencia, S. 45-51.
- Fuchs, G.; Andrienko, N.; Andrienko, G.; Bothe, S.; Stange, H. (2013): Tracing the German centennial flood in the stream of tweets. In: Pfoser, D.; Voisard, A. (Eds.): GEOGROWD 2013: Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information. ACM, New York, S. 31-38.
- Goodchild, M. F. (2007): Citizens as sensors: The world of volunteered geography. In: GeoJournal, 69 (4), S. 211-221.
- Gregorio, F. D.; Varrazzo, D. (2018): psychopg2. <http://initd.org/psycopg/docs/>, Stand 09/2018, Zugriff 03/2019.
- Hahmann, S.; Purves, R.; Burghardt, D. (2014): Twitter location (sometimes) matters: Exploring the relationship between georeferenced tweet content and nearby feature classes. In: Journal of Spatial Information Science, 9, S. 1-36.
- Haklay, M. (2013): Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: Sui, D.; Elwood, S.; Goodchild, M. (Eds.): Crowdsourcing Geographic Knowledge. Springer Netherlands, Dordrecht, S. 105-122.
- Hecht, B.; Hong, L.; Suh, B.; Chi, E. H. (2011): Tweets from Justin Bieber's Heart: The Dynamics of the "Location" Field in User Profiles. In: Tan, D.; Fitzpatrick, G.; Gutwin, C.; Begale, B.; Kellog, W. A. (Eds.): Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, S. 237-246.
- Hendrickson, S.; Kolb, J.; Lehman, B.; Montague, J. (2015): Trend Detection in Social Data. <https://cdn.cms-twigitalassets.com/content/dam/developer-twitter/pdfs-and-files/Trend-Detection.pdf>, Zugriff 08/2020.
- Huang, C.-Y.; Tong, H.; He, J.; Maciejewski, R. (2019): Location Prediction for Tweets. In: Frontiers in Big Data, 2:5. <https://www.frontiersin.org/articles/10.3389/fdata.2019.00005/pdf>.
- Hübner, S.; Vettermann, F. (2016): Erstellung eines Geodatenportals zu den Klein- und Kleinstgewässern in Rostock. In: Bill, R.; Zehner, M.; Golnik, A.; Lerche, T.; Schröder, J.; Seip, S. (Hrsg.): Geoinformation im Alltag: Nutzen und neue Anforderungen. GITO, Berlin, S. 77-83.
- Kim, Y. (2014): Convolutional Neural Networks for Sentence Classification. In: Moschitti, A.; Pang, B.; Daelemans, W. (Eds.): Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Stroudsburg, PA, S. 1746-1751.
- Krüger, R.; Han, Q.; Ivanov, N.; Mahtal, S.; Thom, D.; Pfister, H.; Ertl, T. (2019): Bird's-Eye: Large-Scale Visual Analytics of City Dynamics using Social Location Data. In: Computer Graphics Forum, 38 (3), S. 595-607.
- Lau, J. H.; Chi, L.; Tran, K.-N.; Cohn, T. (2017): End-to-end Network for Twitter Geolocation Prediction and Hashing. In: Liu, C.-H.; Nakov, P.; Xue, N. (Eds.): Proceedings of the IJCNLP 2017. Asian Federation of Natural Language Processing, Taipei.
- Li, R.; Lei, K. H.; Khadiwala, R.; Chang, K. C.-C. (2012): TEDAS: A Twitter-based Event Detection and Analysis System. In: IEEE 28th International Conference on Data Engineering, ICDE 2012, 1. – 5.4.2012, Arlington, VA, S. 1273-1276.
- Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. (2015): Social Sensing: A New Approach to Understanding Our Socioeconomic Environments. In: Annals of the Association of American Geographers, 105 (3), S. 512-530.
- Longley, P. A.; Adnan, M. (2015): Geo-temporal Twitter demographics. In: International Journal of Geographical Information Science, 30 (2), S. 369-389.
- McGrath, R. (2018): twython. <https://twython.readthedocs.io/en/latest/>, Stand 12/2018, Zugriff 03/2019.
- Mehl, D.; Vettermann, F.; Hoffmann, T. G.; Bill, R. (2017): Präferenzen für die Entwicklung kleiner urbaner Gewässer und Feuchtgebiete: Ergeb-

nisse einer Online-Befragung. In: KW – Korrespondenz Wasserwirtschaft, 6, S. 340-346.

Mikolov, T.; Kombrink, S.; Burget, L.; Cernocky, J.; Khudanpur, S. (2011): Extensions of recurrent neural network language model. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 22. – 27.05.2011, Prague Congress Center, Prag. IEEE, Piscataway, NJ, S. 5528-5531.

Miller, H. J.; Goodchild, M. F. (2015): Data-driven geography. In: GeoJournal, 80 (4), S. 449-461.

Nordkurier (2018): Hanse Sail wird wegen Sturm-Warnung unterbrochen. <https://www.nordkurier.de/mecklenburg-vorpommern/hanse-sail-wegen-sturm-warnung-unterbrochen-0932813008.html>, Stand 08/2019, Zugriff 11/2019.

Ostseezeitung (2018): Nach Unfall auf der Sail: „Mir“ sagt alle Ausfahrten ab. <https://www.ostsee-zeitung.de/Thema/H/Hanse-Sail/Nach-Unfall-auf-der-Sail-Mir-sagt-alle-Ausfahrten-ab>, Stand 08/2018, Zugriff 11/2019.

Recchia, G.; Louwse, M. (2013): A Comparison of String Similarity Measures for Toponym Matching. In: Scheider, S.; Adams, B.; Janowicz, K.; Vasardani, M.; Winter, S. (Eds.): ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems: ACM SIGSPATIAL International Workshop on Computational Models of Place. Association for Computing Machinery Inc., Orlando, S. 54-61.

Resch, B. (2017): Nutzergenerierte Daten für Entscheidungsunterstützung in naher Echtzeit. München.

Röder, M.; Both, A.; Hinneburg, A. (2015): Exploring the Space of Topic Coherence Measures. In: Cheng, X.; Li, H.; Gabrilovich, E.; Tang, J. (Eds.): WSDM'15: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 31.01. – 06.02.2015, Shanghai. ACM Association for Computing Machinery, New York, NY, S. 399-408.

Scheffler, T. (2014): A German Twitter Snapshot. In: Calzolari, N.; Choukri, K.; Declerck, T.; Loftson, H.; Maegaard, B.; Mariani, J.; Moreno, A.; Odijk, J.; Piperidis, S. (Eds.): IREC 2014, Ninth International Conference on Language Resources and Evaluation, 26. – 31.05.2014, Reykjavik. European Language Resources Asso-

ciation (ELRA), Reykjavik, S. 2284-2289.

Snyder, L. S.; Karimzadeh, M.; Chen, R.; Ebert, D. S. (2019): City-level Geolocation of Tweets for Realtime Visual Analytics. In: Bouros, P.; Dasu, T.; Kanza, Y.; Renz, M. (Eds.): LocalRec '19: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-based Recommendations, Geosocial Networks and Geoadvertising. Association for Computing Machinery, New York, S. 85-88.

Streich, B. (2014): Subversive Stadtplanung. Springer Fachmedien, Wiesbaden.

Süddeutsche Zeitung (2019): Positives Fazit der 29. Hanse Sail: Großer Besucherzuspruch. <https://www.sueddeutsche.de/kultur/feste-rostock-positives-fazit-der-29-hanse-sail-grosser-besucher-zuspruch-dpa-urn-newsml-dpa-com-20090101-190810-99-417000>, Stand 08/2019, Zugriff 11/2019.

Sundermeyer, M.; Schlüter, R.; Ney, H. (2013): Lstm neural networks for language modeling. In: ISCA (Ed.): 13th Annual Conference of the International Speech Communication Association 2012 (INTERSPEECH 2012), 9. – 13.09.2012. Portland, Oregon. Curran, Red Hook, NY, S. 194-197.

Terpstra, T.; Vries, A.; Stronkman, R.; Pradies, G. L. (2012): Towards a realtime Twitter analysis during crises for operational crisis management. In: Rothkrantz, L.; Ristvej, J.; Franco, Z. (Eds.): IS-CRAM 2012 Conference Proceedings: 9th International Conference on Information Systems for Crisis Response and Management. Simon Fraser University, Vancouver.

Tourismuszentrale Rostock & Warnemünde (2018): Hanse Sail 2018: Schiffe, Land- und Bühnenprogramm. <https://www.rostock-heute.de/hanse-sail-2018-rostock/101949>, Stand 08/2018.

Tränckner, J.; Mehl, D.; Thiele, V. (2017): Integrale Gewässerentwicklung auf der Ebene einer Großstadt: Das Projekt KOGGE. In: KW – Korrespondenz Wasserwirtschaft, 4, S. 209-215.

Traynor, D.; Curran, K. (2013): Location-Based Social Networks. In: Lee, I. (Ed.): Mobile Services Industries, Technologies, and Applications in the Global Economy. Information Science Reference, Hershey, S. 243-253.

Twitter Inc. (2015): Annual Report 2015.

Twitter Inc. (2019): Häufig gestellte Fragen zu Trends auf Twitter. <https://help.twitter.com/de/using-twitter/twitter-trending-faqs>, Stand 2019, Zugriff 11/2019.

Vettermann, F. (2019): Extraktion und Auswertung von Geodaten aus Sozialen Netzwerken als Element der Bürgerbeteiligung in kommunalen Belangen der Hansestadt Rostock. Veröffentlichungen der DGK, Reihe C, 836, S. 1-180. https://dgk.badw.de/fileadmin/user_upload/Files/DGK/docs/c-836.pdf.

Vettermann, F.; Bill, R. (2019): Rostock auf Twitter: Geolokalisation und Auswertung von Tweets mit Bezug zur Hanse- und Universitätsstadt Rostock. In: AGIT – Journal für Angewandte Geoinformatik, 5-2019, S. 292-304.

Vettermann, F.; Seip, C.; Bill, R. (2017a): Die Hanse Sail 2016 auf Twitter: Nutzung von Geolokalisation in Sozialen Netzwerken im kleinräumigen Maßstab. In: Bill, R.; Zehner, M.; Golnik, A.; Lerche, T.; Schröder, J.; Seip, S. (Hrsg.): Mit Geoinformation Planen. GITO, Berlin, S. 123-131.

Vettermann, F.; Seip, C.; Bill, R. (2017b): Using Twitter for Geolocation Purposes during the Hanse Sail 2016 in Rostock. In: Oj Jacques, B.; Hitzelberger, P.; Naumann, S.; Wohlgemuth, V. (Eds.): From Science to Society: New Trends in Environmental Informatics. Progress in IS. Springer, Cham, S. 171-180.

Vettermann, F.; Weinzierl, T.; Bill, R. (2018): Monitoring von Twitter-Nachrichten zur raumzeitlichen und thematischen Analyse in der Hansestadt Rostock: Der Twittermonitor Rostock. In: gis. Science, 1/2018, S. 1-9.

Wold, H. M.; Vikre, L.; Gulla, J. A.; Özgöbek, Ö.; Su, X. (2016): Twitter Topic Modeling for Breaking News Detection. In: Majchrzak, T. A.; Traverso, P.; Monfort, V.; Krempels, K.-H. (Eds.): Proceedings of the 12th International Conference on Web Information Systems and Technologies. SCITEPRESS – Science and Technology Publications Lda, Setúbal, S. 211-218.