# HUSO 2017

The Third International Conference on Human and Social Analytics

ISBN: 978-1-61208-578-4

July 23 - 27, 2017

Nice, France

**HUSO 2017 Editors**

Els Lefever, LT3/Ghent University, Belgium

Jan Ole Berndt, Trier University, Germany

Dennis J. Folds, Georgia Tech Research Institute, USA

# HUSO 2017

# Foreword

The Third International Conference on Human and Social Analytics (HUSO 2017), held between July 23 - 27, 2017 - Nice, France continued the inaugural event bridging the concepts and the communities dealing with emotion-driven systems, sentiment analysis, personalized analytics, social human analytics, and social computing.

The recent development of social networks, numerous ad hoc interest-based formed virtual communities, and citizen-driven institutional initiatives raise a series of new challenges in considering human behavior, both on personal and collective contexts.

There is a great possibility to capture particular and general public opinions, allowing individual or collective behavioral predictions. This also raises many challenges, on capturing, interpreting and representing such behavioral aspects. While scientific communities face now new paradigms, such as designing emotion-driven systems, dynamicity of social networks, and integrating personalized data with public knowledge bases, the business world looks for marketing and financial prediction.

We take here the opportunity to warmly thank all the members of the HUSO 2017 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to HUSO 2017. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the HUSO 2017 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that HUSO 2017 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the area of human and social analytics.

We are convinced that the participants found the event useful and communications very open. We also hope that Nice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

**HUSO 2017 Chairs:**

**HUSO Steering Committee**

Els Lefever, Ghent University, Belgium
Dennis J. Folds, Georgia Tech Research Institute, USA
Nitin Agarwal, University of Arkansas at Little Rock, USA
Chen Ding, Ryerson University, Canada

Liz Bacon, University of Greenwich, UK
Jan Ole Berndt, University of Trier, Germany
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Meena Jha, CQUniversity Sydney, Australia
Xiaolong Jin, 靳小龙, Chinese Academy of Sciences, China

**HUSO Industry/Research Advisory Committee**

Baden Hughes, Glentworth Consulting, Australia
Shams Zawoad, Visa Inc., USA
Yi Shan, SDE II, Electronic Arts, Seattle, USA
Fan Yang, eBay Inc., USA
Massimiliano Zanin, The Innaxis Foundation & Research Institute, Madrid, Spain
Zhiwen Fang, Microsoft, USA

# HUSO 2017

# Committee

**HUSO Steering Committee**

Els Lefever, Ghent University, Belgium
Dennis J. Folds, Georgia Tech Research Institute, USA
Nitin Agarwal, University of Arkansas at Little Rock, USA
Chen Ding, Ryerson University, Canada
Liz Bacon, University of Greenwich, UK
Jan Ole Berndt, University of Trier, Germany
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Meena Jha, CQUniversity Sydney, Australia
Xiaolong Jin, 靳小龙, Chinese Academy of Sciences, China

**HUSO Industry/Research Advisory Committee**

Baden Hughes, Glentworth Consulting, Australia
Shams Zawoad, Visa Inc., USA
Yi Shan, SDE II, Electronic Arts, Seattle, USA
Fan Yang, eBay Inc., USA
Massimiliano Zanin, The Innaxis Foundation & Research Institute, Madrid, Spain
Zhiwen Fang, Microsoft, USA

**HUSO 2017 Technical Program Committee**

Paul Abbiati, Founding Fellow of ELI - The European Law Institute, Austria
Ifeoma Adaji, University of Saskatchewan, Canada
Nitin Agarwal, University of Arkansas at Little Rock, USA
Rodrigo Agerri, University of the Basque Country (UPV/*EHU*), Spain
Harry Agius, Brunel University London, UK
Hafizi Muhamad Ali, Yanbu University College, Saudi Arabia
Alessio Antonini, Università degli Studi di Torino, Italy
Saïd Assar, Telecom Ecole de Management, France
Liz Bacon, University of Greenwich, UK
Panagiotis D. Bamidis, Aristotle University of Thessaloniki, Greece
Stefano Berretti, University of Florence, Italy
Shreyansh Bhatt, Kno.e.sis | Wright State University, USA
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Stefan Bratosin, Université Paul Valéry Montpellier 3, France
Luca Cagliero, Politecnico di Torino, Italy
María Luisa Carrió Pastor, Universitat Politècnica de València, Spain
Mario Cataldi, University of Paris 8, France
Chin-Chen Chang, Feng Chia University, Taichung, Taiwan

Bo-Hao Chen, Yuan Ze University, Taiwan
Dickson Chiu, The University of Hong Kong, Hong Kong
Yung Ryn (Elisha) Choe, Sandia National Laboratories, Livermore, USA
Aresh Dadlani, Gwangju Institute of Science and Technology (GIST), South Korea
Orphée De Clercq, LT3 | Ghent University, Belgium
Sara de Freitas, Murdoch University, Australia
Ibrahim Delibalta, Intel Corporation, USA / Koc University, Istanbul, Turkey
Engin Deniz Eris, Dokuz Eylül University, Turkey
Luigi Di Caro, University of Turin, Italy
Chen Ding, Ryerson University, Canada
Joyce El Haddad, Université Paris Dauphine, France
Samy El-Tawab, James Madison University, USA
Zhiwen Fang, Microsoft, USA
Silvia Florea, Lucian Blaga University of Sibiu, Romania
Dennis J. Folds, Georgia Tech Research Institute, USA
Paolo Garza, Politecnico di Torino, Italy
Feng George Yu, Youngstown State University, USA
Denis Gracanin, Virginia Tech, USA
William Grosky, University of Michigan-Dearborn, USA
Diofantos G. Hadjimitsis, Cyprus University of Technology, Cyprus
Malka N. Halgamuge, University of Melbourne, Australia
Ragib Hasan, University of Alabama at Birmingham, USA
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Baden Hughes, Glentworth Consulting, Australia
Emilio Insfran, Universitat Politecnica de Valencia, Spain
Roberto Interdonato, Uppsala University, Sweden
Jaroslaw Jankowski, Wrocław University of Technology, Poland
Meena Jha, CQUniversity Sydney, Australia
Minhao Jiang, The Hong Kong University of Science and Technology, Hong Kong
Xiaolong Jin, Institute of Computing Technology (ICT) | Chinese Academy of Sciences (CAS), China
Mirko Kämpf, Cloudera, Inc., USA
Selvi Kannan, College of Business | Victoria University, Australia
Hassan A. Karimi, University of Pittsburgh, USA
Konstantin Kuzmin, Rensselaer Polytechnic Institute (RPI), USA
Georgios Lappas, Western Macedonia University of Applied Sciences, Greece
Els Lefever, Ghent University, Belgium
Ken Li, Hong Kong Institute of Vocational Education, Hong Kong
Rita Yi Man Li, Hong Kong Shue Yan University, Hong Kong
Mimie (Dimitra) Liotsiou, University of Southampton, UK
Francesco Longo, Università degli Studi di Messina, Italy
Eurico Lopes, Instituto Politécnico de Castelo Branco, Portugal
Yutao Ma, Wuhan University, China
Christos Makris, University of Patras, Greece
Mahshid Marbouti , University of Calgary, Canada
Elvis Mazzoni, University of Bologna, Italy
Fernanda Monteiro Eliott, Vanderbilt University, USA
Jason R.C. Nurse, University of Oxford, UK
Jan Ole Berndt, University of Trier, Germany

## Table of Contents

# Forecasting Civil Strife: An Emerging Methodology

Dipak K. Gupta*, Sathappan Muthiah‡, David Mares† and Naren Ramakrishnan‡
*Professor Emeritus
Department of Political Science
San Diego State University
†Chair for Inter-American Affairs
Institute of the Americas
University of California, San Diego
‡Discovery Analytics Center
Department of Computer Science
Virginia Tech, Arlington, VA

***Abstract*—From the earliest time of recorded scholarship, forecasting civil strife has been the Holy Grail to political theorists. Yet, without actual data and ability to conduct empirical analyses, until the 1960's such analyses were no more than speculation. The advent of high-speed computing along with collection of data on civil unrest allowed political scientists to empirically test their hypotheses. Yet, these analyses did not result in short term prediction due to the lack of real time data. Today the rise of social media has witnessed a radically different methodology in how we can understand, monitor, and forecast incidents of social strife in real time. This emerging methodology, however, requires a multi-disciplinary effort no one even contemplated until recently. This paper presents results of forecasting events of politically motivated violence based on monitoring open source information (Twitter, blogs, newspaper articles) in 10 Latin American countries by a multi-university, multi-disciplinary team of academics, supported by a grant from the Intelligence Advanced Research Projects Activity (IARPA).**

***Keywords–Forecasting, Civil Strife, Social Media, Multidisciplinary Methodology***

## I. INTRODUCTION

Generally speaking, social scientists are reluctant to forecast incidents of civil unrest. The open system, within which these outcomes are derived, has been considered far too complex to forecast anything more than the direction of trends or the increased likelihood of an event occurring. In fact, the noted social scientist Paul Collier [1] flatly admits concerning his own model: "More fundamentally, our model (of political conflict) cannot be used for prediction, it cannot tell you whether Sierra Leone will have another civil war next year. That depends on a myriad of short-term events." As a result, the efforts of social scientists have largely been confined to the understanding of the determinant variables using various types of causal models or by using time-series projections. From a public policy standpoint, the importance of forecasting imminent political upheavals is undeniable. The urgency of developing such predictive capabilities has been further precipitated by the spread of social media and the Internet. The prime example of the potential of social media to mobilize civil strife was amply demonstrated during the sudden uprising, dubbed, the "Arab Spring." Rise of social media has been a double-edged sword; along with its ability to mobilize the disgruntled, it has accorded us a new avenue through which we can understand society in a way that was impossible in the past. Our ability to understand and forecast civil strife has been vastly expanded because of two primary reasons. First, social media, especially in democratic nations, allows individuals to express their opinions freely. As a result, when monitored properly, social media allows us to comprehend societal trends in a way that even the best-constructed opinion surveys are not able to capture. Second, thanks to the recent advancements in a number of related academic disciplines as well as computational capabilities, we are witnessing a revolutionary change in how we understand and then predict various societal events [2] [3], such as spreading chaos in the financial market [4] [5], spread of infectious diseases [6]–[9], etc. These efforts require a multi-disciplinary focus that was until recently largely absent.

Our multidisciplinary research effort, which brings in expertise from wide-ranging disciplines across social sciences, linguistics, geographic information systems, and computer science, concentrates on "Civil Strife", by which we mean mass movements, such as protest movements and riots and other acts of collective rebellion and not on terrorism or the so-called "lone wolf" attacks, plotted and carried out by a small group or an isolated individual.

For our study, we chose 10 large Latin American countries. These countries offer a unique combination of mostly functioning democracies with a long vibrant tradition of public discourse of political issues along with excellent penetration of Internet communication.

Our article is divided in four broad parts. The first part introduces the problems of forecasting from a social science perspective and provides a theoretical basis of human motivation for participating in mass movements (Section I and Section II). The second part provides the foundation of our subsequent analyses (sections III and IV). The third part (Section V) explains the index of accuracy of our research effort. The final part (Section VI) of the paper discusses ways of moving our research forward.

## II. THEORETICAL BASIS OF FORECASTING THROUGH SOCIAL MEDIA

Social media is a noisy medium, where individuals and groups participate using messages that are confusing in language, terminologies, and expressions, made even more complicated by their temporal malleability. Therefore, before we begin to make sense of this vast and chaotic linguistic landscape, we need to put our analysis within a broad theoretical framework.

As scholars looked deeper into the process by which individual participants are moved into creating collective actions, Olson [10] raised a theoretical concern. When it comes to volunteering for a political cause, there is an inevitable inertia of what is known in social sciences as the "collective action" problem, where a rational actor asks the inevitable question, "why me?" The answer to this question may be found in the work of social psychologists, starting with the path breaking work of Tajfel and his associates [11] [12]. Their work, widely accepted as "social identity" theory argues that, contrary to the economic assumption of human behavior that equates self-interest with human rationality [13], individuals are also motivated to voluntarily participate in collective actions out of their community concerns. Group identity as a political force, however, does not develop spontaneously [14]. For that, the society needs to have leaders or "political entrepreneurs." These leaders need to clearly define the perimeters of the "us" and "them," by instilling in the minds of the adherents, the most primal emotion of all: fear [15].

We can forecast future actions in social media by searching for markers of extant grievances ("injustice", "deprivation", "inequality", etc.), the boundaries of the in-group (viz., "students", "workers", "farmers", "indigenous people", etc. along with the names of the "heroes" of the community) the out-group (viz., "politicians", "thieves", "oppressors", "police" etc., along with names of the "villains" of the movement). Finally, we can also track the actions suggested by the leaders (viz., "protest", "march", "destroy", etc.). Based on this theoretical framework, we developed a library of keywords, reflecting existing grievance, the leaders' framing of collective identity, and prodding of actions against the offending party. The results of tracking such keywords in social media lead us to our forecasts of impending civil strife. For our forecasting, we found a dichotomy between planned and unplanned events to be useful. We search for "planned events" by searching for news or announcements by various organizations for upcoming protest events. Figure 1 illustrates the theoretical framework, where political entrepreneurs take the prevailing grievances and frame those in terms of "us" and "them," provide an action plan to bring about events of political demonstrations or riots. By tracking this process, we predict incidents of civil strife in a country.

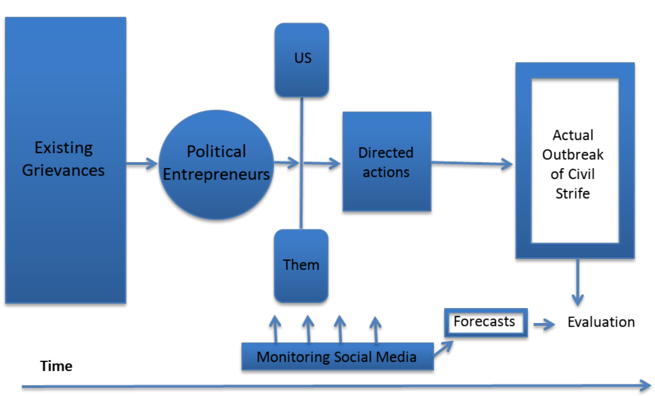


Figure 1. Theoretical Framework of Forecasting Civil Strife

## III. Features of Our Study

Our study offers a number of important features:

1) On the Internet there is information that is public ("blasts" to world) and private (communication between individuals or groups, otherwise protected by law). We strictly use only public information obtained from open sources, such as newspapers, blogs, and public areas of social media such as Facebook and Twitter. In other words, none of our data are private or classified.
2) The events of civil strife were classified by strict definitions of who participates (general population, labor unions ...), issues (political, environmental ...), geographic location, and time of event. The *who* part of an event is described by the categories defined in the "social" sector of the TABARI [16] event coding system. The issues or the *why* part of an event is classified into 7 categories. Apart from encoding the *who* and *why* part of an event, our study also strives to identify if a civil strife will turn violent or not. An event is deemed violent if there is significant damage to property or if there is any act of violence directly associated with the event that results in injuries to anyone involved. The geographic location entry of an event description uses a 3-level typology with the country at the top, province or state (admin level 1) in the middle and city at the last level. The geonames gazetteer [17] is used as the reference standard for the location specification of an event. Finally, the time of an event is encoded with date level accuracy.
3) The actual post-facto events were recorded from the local, national, and international newspapers and codified by a third party (in this case, MITRE Corporation) and were published in IARPA's Gold Standard Report (GSR). The list of newspapers were identified using the rankings as provided by 4 International Media & Newspapers [18] and with subject matter expert input.
4) Forecasts had to be submitted in real time and at the end of every month would be matched against the GSR by MITRE Corporation and rated for accuracy as well as lead time. False positives and false negatives were also tracked as were the confidence in the forecasts.
5) Ultimately, the delivery of warnings had to be fully automated, without a "human in the loop". Humans may be involved in the development and training phases of modeling, but the final warnings must be fully machine generated and automatically submitted without any interference (in form of filtering or guiding) from subject matter experts.For illustrative purposes, we have also presented a sample of a hypothetical warning in Figure 2.
6) Every forecast should have an audit trail for IARPA to trace it back to the causal factors.

## IV. Algorithms for Forecasting

In this section, we provide details of how open source data is harvested and enriched by our system and finally how this

 ISBN: 978-1-61208-578-4

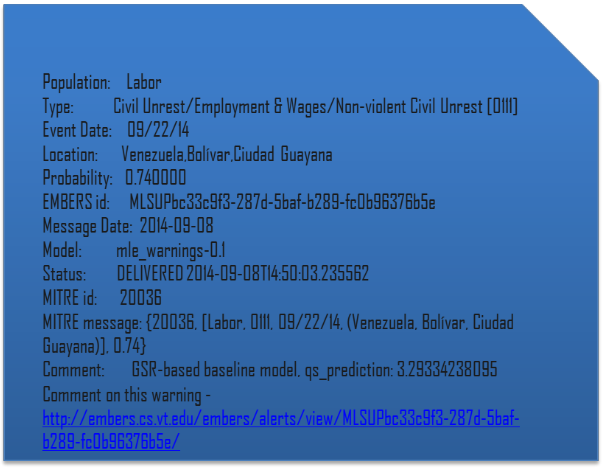


Figure 2. An example of computer generated warning

enriched data is used in various models to produce forecasts or alerts.


Figure 3. System Architecture

### A. Data Collection and Enrichment

Our work begins with searching the vast open source (public) areas of the Internet by following the keywords seeded by social scientists (the subject matter experts). Figure 3 explains our system architecture. These searches scoop up a significant quantity of linguistic information, much of which are unrelated to our objective. These data are stored in a "vat" and "enriched" by a Natural Language Processing(NLP) pipeline (includes language identification, tokenization, lemmatization, named entity recognition and Part-of-Speech tagging) to make the necessary connections between words and their sought meaning. The Basis technologies RLP suite [19] is used for all the NLP tasks. The textual data is then passed through a temporal normalization system like Heideltime [20] which converts any relative temporal expressions such as *today, tomorrow, a week ago* to absolute time based on the article/tweets publish time. It is not enough, however, to understand the underlying meaning of the collected keywords; for forecasting, we need to also know the geographic points of origin. A geographic information system allows us to geo-locate these conversations onto maps. We use different geocoding systems for different types of data sources. For example, twitter geocoding is achieved by looking at different parts of a tweet in an orderly fashion starting from the least available source, viz. geotags, followed by Twitter places field, the text fields contained in user profile (location, description) and finally the tweet text itself to find mentions of relevant locations. For news articles and blog posts, we develop a probabilistic reasoning engine using Probabilistic Soft Logic (PSL) [21] to identify which among the multiple location names mentioned within the text(such as the location of reporting, the incident location, etc.) is the main geographic focus of the article. These data points are now ready for statistical analyses, yielding forecasts.

### B. Data Modeling

In this section we introduce briefly some of the main prediction models that work on different real-time datasets to produce alerts and a fusion engine that combines these alerts to achieve high quality and performance.

Protests in many countries are regular occurrences (for example Mexico has an average of 245.95 events per month in the GSR during 2013-2014). A baseline model that produces forecasts using only historical occurrences of events and no real-time data performs reasonably well and can be beat only in situations where there is an abnormal spike (or dip) in number of events The baseline model is trained using the ground truth event data from the previous three months to make forecasts for the current month. The baseline model estimates the expected number of events for the current month, for a given location, reason (or issue) and population (the *who* part) as the average of the number of such events in the training period. Finally, the time information (expected date of occurrence of event) is obtained by performing a random draw from a uniform distribution. We have conducted a maximum entropy analysis wherein we evaluated the baseline model's ability to forecast surprising events (i.e., events whose frequency falls outside the norms of historical data) and found that EMBERS consistently beats the baseline model [22]. EMBERS uses a suite of six forecasting algorithms that posit different approaches to modeling civil unrest.

The planned protest model aims to identify incidents of organized and preannounced protests from news and Twitter using language-processing techniques [23]. It searches for occurrences of phrases denoting "calls for protest" (like "planar protesta", "Announce Strike", "Manhã de mobilização",.etc) in news and Twitter along with the mention of a future date to produce alerts. The phrase list is developed in a semi-automatic manner using template based matching techniques on about 1 year worth of data (twitter, news articles and blog posts) and has mostly remained static barring a few changes for the entire duration of our project.

A second model uses spatial scan statistics to identify geo-located clusters of tweets enriched with a defined vo-

 ISBN: 978-1-61208-578-4 

cabulary of 726 keywords [24].The goal of the model is to identify anomalous spatial regions based on Poisson mixtures. A fast linear time subset scan [25] is applied to identify the anomalous regions and each such region is scored using p-values computed by Monte Carlo simulations. The keyword filtered twitter stream is split into day level chunks and the fast subset scan is run separately for each day. A spatial cluster on any day is considered to be an continuation of a previous day cluster if the two clusters have a jaccard similarity (between the keywords present in each cluster) is greater than a threshold. In this manner the growth of a cluster both in terms of size and density is tracked over time and these characteristics are then used to issue a forecast for a given spatial region. The spatial scan model is trained using 6 months of twitter data. In comparison to the planned protest model, this model can track the development of a protest from its birth.However this models performance is limited by the amount of coverage/popularity the issue of a protest achieves on social media as not all protests are covered by social media like twitter. Also since social media is more prone to rumors it can lead to alerts being generated falsely.

The cascade regression model recognizes situations where social media, such as Twitter, is utilized as the staging ground for galvanizing support for protests via online recruitment to the underlying causes [26]. The model studies activity cascades in twitter to understand spread of influence and information and tries to forecast date of event based on this information spread besides identifying a critical subset of users responsible for the formation and survival of the activity cascade. The model analyzes over 353 million tweets over a 1.5 year period. Each tweet contains at least 3 keywords from a dictionary of over 900 keywords in 3 languages related to civil unrest. This ensures the activity cascades are relevant to our topic of interest. Next, for making forecasts a regression model (LASSO) using a feature set based on structural properties of cascades like size of cascade, number of participants, duration of cascades, change in the number of participants and tweets, average growth rate of tweets ,.etc, is used to predict the probability of a civil unrest event in a given day.

The dynamic query expansion model is similar to the spatial scan model but aims to learn new emerging keywords, unlike the static vocabulary used by the spatial scan model [27]. In 2013, for instance, there were a series of protests in Venezuela due to a shortage of toilet paper, a novel circumstance that was uncovered using this model. The dynamic query expansion model starts with a very small set of seed keywords (like protest, march, demonstration etc.) and iterates through the data identifying semantically similar and co-occurring terms. The model repetitively sweeps through the data, learning a larger set of relevant keywords at each iteration, unless it converges. This dynamic learning of new relevant keywords helps the model identify novel/unusual circumstances of protest.

Both the cascade regression model and dynamic query expansion suffer from the same disadvantages as the spatial scan model and is trained on at least 6 months worth of historical data.

The volume-based LASSO model uses every possible data source in our study (news, tweets, blogs, economic indicators, TOR, and smiles) to forecast the imminence of protests in the next day or two [28]. The volume-based LASSO model provides insights about the underlying social dynamics in different countries by identifying predictive features that are tied to unrest. The model also is capable of identifying the value of different data source in predicting an unrest and is more robust to changes in individual data characteristics as compared to the previously mentioned models. The main disadvantages of such a model is low recall and lead-time. Also, this model works better at country and state levels as opposed to city level (not all datasets are available at city level granularity)

Finally, the MLE(Maximum Likelihood Estimate) model aims to identify regularities in the GSR and provides a baseline performance level [29]. Each of these models is tuned for high precision, and their fusion aims to achieve high recall.

All the above-mentioned models produce alerts independent of each other with each model tuned for high precision. The fusion model then aims at combining the different predictions from these individual models to achieve high quality forecasts with tunable precision and recall. The goal of the fusion model is to (i)identify duplicate alerts (alerts with same location, date, type and population) as models share data sources and thus their hypothesis space overlap, (ii) fill-in missing values if any in an alert (for example certain forecasts of the volume based model are only at country level and it is the fusion engines responsibility to add city information) , (iii) re-write warning fields if necessary as it is possible for a model to issue alert for an improbable combination of <*date,location,type,population*> due to noisy data and finally, (iv) balance recall and quality. The recall-quality trade-off is achieved by first building a random forest regression model to predict the expected quality score of an alert and then a threshold set on this expected quality score can be used to tune the precision and recall (with a lower bound on quality) of our overall system.

## V. Accuracy of Forecasts

The forecasts for a given month are evaluated by MITRE at the end of every month using the Gold Standard Report (GSR). The forecasts are matched against the GSR using the Hungarian [30] bipartite matching algorithm. At the end of the bipartite matching set of alert-event match pairs are obtained along with the list of unmatched events and unmatched alerts. Each alert-event match pair is assigned a quality score out of 4. The problem of measuring forecasting accuracy is that each forecast is a multi-dimensional phenomenon. Thus, a specific forecast must match the time and geographic place where it was supposed to have taken place and it should also correspond to the type of protest and its specific cause. Thus, the quality score is a composite measure of four components. These are: location score, date score, event-type score and event-population score. Precision refers to the fraction of alerts that got matched to a true event, whereas recall refers the fraction of true events (GSR) that got matched to an alert (i.e., were accurately forecasted). The average lead-time of the system of was 9.76 days i.e., on an average our system produced an alert for a civil unrest event 9.76 days in advance of the first report of the event.

Our ability to forecast events in Latin America was put to test by a sudden explosion of public anger in Brazil during mid-2013 and in Venezuela in early 2014. Table I provides the performance metrics for the two countries during these

TABLE I. COMPARISON OF ACCURACY FOR BRAZIL AND VENEZUELA

| *Country* | *Period* | *Quality* | *Precision* | *Recall* | *Lead-Time* |
|---|---|---|---|---|---|
| Brazil | May'13-Aug'13 | 3.44 | 0.53 | 0.69 | 7.04 |
| Venezuela | Feb'14 - March'14 | 3.66 | 0.91 | 0.50 | 2.25 |

periods. Similar to the so-called "Arab Spring," where three years earlier the self-immolation by an obscure fruit vendor in Tunisia's capital city touched off a cascading wave of protests engulfing nearly the entire Arab world, a sudden spread of protests inundated Brazil. Clearly, these were not protests that were planned weeks in advance. Therefore, our models had to quickly adapt to the rapidly changing political landscape with government forces interacting with the protesters, sometime quelling, other times adding fuel to the fire through their highhanded reactions. We have presented our forecasts in Figure 4. The demonstrations originally started in protest of rising bus prices in June. We have also presented our results of Venezuelan protests of February 2014 in Figure 5. The map of geolocation for the Venezuelan protests is shown in Figure 6.

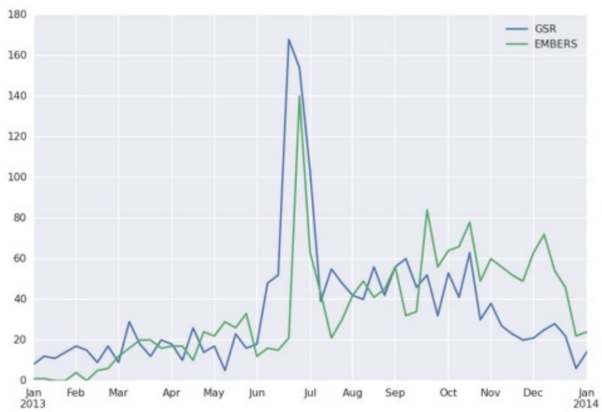


Figure 4. Forecasting performance during Brazil Spring

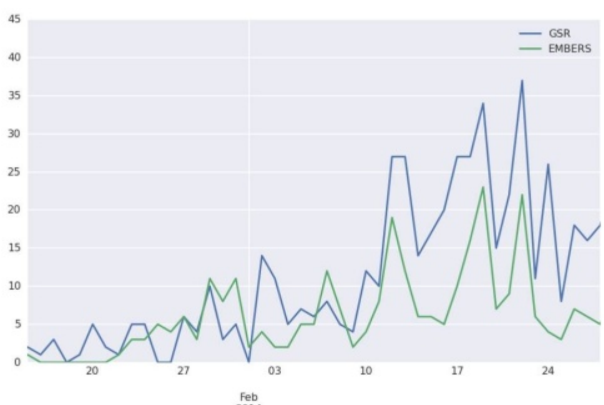


Figure 5. Forecasting performance in Venezuela 2014

## VI. DISCUSSION: A BRAVE NEW WORLD OF FORECASTING AND CONTROL?

The rise of social media has allowed us to break down the barriers of geographic space and to create virtual communities


Figure 6. Geolocation of Venezuelan Protests

of like-minded people. This trend has simultaneously united and divided people all over the world. Our current effort aims at detecting these trends to forecast incidents of civil strife in real time. Similar to any new innovation, this emerging methodology has its obvious downside; it carries the risk of being a tool of state oppression. Authoritarian regimes all over the world are trying to find ways to control their citizens by manipulating Internet conversations, particularly in the aftermath of the "Arab Spring." Given the newness of the technology, the ethical implications of all of these issues are still evolving. However, there are some important considerations related to forecasting acts of civil strife. First, in light of recent events there is a heightened awareness of privacy issues surrounding developing surveillance capabilities, even when they involve publicly available information.Yet, our project does not use any information that is not publicly available. Our project does not use any information that is not publicly available. Our project has developed a completely automated system based on 10 Latin American nations that begins with monitoring social media and ends with generating warnings in one smooth loop.Second, the genie is out of the bottle. As new modes of social media spring up, their use spread throughout the world and algorithms are perfected to monitor the conversations-whether for public policies or for private commercial gains - there is simply no effective way of stopping these monitoring efforts.

There is little doubt that we are witnessing a brand new world of information processing and trans-discipline inquiries. Although we can clearly see that we have developed the capability for real-time forecasting of incidents of civil strife, we must realize that our efforts at forecasting events is only in the short term. As for the long-term course of a mass movement, we must agree with Collier [1] ; it is impossible to conceive of any algorithm that would be able to do an effective job. Finally, our effort was strictly focused on forecasting and not about finding causal relationships that create political instability. We must use this methodology to gain a deeper insight into this aspect of academic inquiry.

## ACKNOWLEDGEMENT

Our project (EMBERS: Early Model-Based Event Recognition using Surrogates) is supported by the Intelligence Advanced Research Projects Activity (IARPA) via DoI/NBC con-

 ISBN: 978-1-61208-578-4 

tract number D12PC000337, the US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the US Government.

## REFERENCES


[1] P. Collier, "Bottom billion," The Wiley-Blackwell Encyclopedia of Globalization, 2007.

[2] L. An, M.-H. Tsou, S. E. Crook, Y. Chun, B. Spitzberg, J. M. Gawron, and D. K. Gupta, "Space–time analysis: Concepts, quantitative methods, and future directions," Annals of the Association of American Geographers, vol. 105, no. 5, 2015, pp. 891–914.

[3] L. An, M.-H. Tsou, B. H. Spitzberg, D. K. Gupta, and J. M. Gawron, "Latent trajectory models for space-time analysis: An application in deciphering spatial panel data," Geographical Analysis, vol. 48, no. 3, 2016, pp. 314–336.

[4] H. Mao, S. Counts, and J. Bollen, "Predicting financial markets: Comparing survey, news, twitter and search engine data," arXiv preprint arXiv:1112.1051, 2011.

[5] E. J. Ruiz, V. Hristidis, C. Castillo, A. Gionis, and A. Jaimes, "Correlating financial time series with micro-blogging activity," in Proceedings of the fifth ACM international conference on Web search and data mining. ACM, 2012, pp. 513–522.

[6] K. Lee, A. Agrawal, and A. Choudhary, "Real-time disease surveillance using twitter data: demonstration on flu and cancer," in Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, 2013, pp. 1474–1477.

[7] H. Achrekar, A. Gandhe, R. Lazarus, S.-H. Yu, and B. Liu, "Predicting flu trends using twitter data," in Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on. IEEE, 2011, pp. 702–707.

[8] A. A. Aslam et al., "The reliability of tweets as a supplementary method of seasonal influenza surveillance," Journal of medical Internet research, vol. 16, no. 11, 2014.

[9] A. C. Nagel et al., "The complex relationship of realspace events and messages in cyberspace: case study of influenza and pertussis using tweets," Journal of medical Internet research, vol. 15, no. 10, 2013.

[10] M. Olson, The logic of collective action. Harvard University Press, 2009, vol. 124.

[11] H. Tajfel and J. Turner, "The social identity theory of inter group behavior in s worchel & wg austin (eds) psychology of intergroup relations," Chicago: Nelson, 1986.

[12] H. Tajfel, "Differentiations between social groups: Studies in intergroup relations," 1978.

[13] A. K. Sen, "Rational fools: A critique of the behavioral foundations of economic theory," Philosophy & Public Affairs, 1977, pp. 317–344.

[14] D. K. Gupta, Understanding terrorism and political violence: The life cycle of birth, growth, transformation, and demise. Routledge, 2008.

[15] J. M. Gawron, D. Gupta, K. Stephens, M.-H. Tsou, B. H. Spitzberg, and L. An, "Using group membership markers for group identification." in ICWSM, 2012.

[16] P. A. Schrodt, "Automated production of high-volume, near-real-time political event data," in American Political Science Association meetings, 2010.

[17] M. Wick and B. Vatant, "The geonames geographical database," Available from World Wide Web: http://geonames. org, 2012.

[18] http://www.4imn.com/, [Online; accessed 2012].

[19] "Basis Technology Rosette Linguistics Platform," http://www.basistech.com/text-analytics/rosette/.

[20] J. Strötgen and M. Gertz, "Heideltime: High quality rule-based extraction and normalization of temporal expressions," in Proceedings of the 5th International Workshop on Semantic Evaluation, ser. SemEval '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 321–324. [Online]. Available: http://dl.acm.org/citation.cfm?id=1859664.1859735

[21] A. Kimmig, S. Bach, M. Broecheler, B. Huang, and L. Getoor, "A short introduction to probabilistic soft logic," in Proceedings of the NIPS Workshop on Probabilistic Programming: Foundations and Applications, 2012, pp. 1–4.

[22] S. Muthiah et al., "Embers at 4 years: Experiences operating an open source indicators forecasting system," in Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 205–214. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939709

[23] S. Muthiah, B. Huang, J. Arredondo, D. Mares, L. Getoor, G. Katz, and N. Ramakrishnan, "Planned protest modeling in news and social media." in AAAI, 2015, pp. 3920–3927.

[24] F. Chen, J. Arredondo, R. P. Khandpur, C.-T. Lu, D. Mares, D. Gupta, and N. Ramakrishnan, "Spatial surrogates to forecast social mobilization and civil unrests," in Position Paper in CCC Workshop on "From GPS and Virtual Globes to Spatial Computing-2012, 2012.

[25] D. B. Neill, "Fast subset scan for spatial pattern detection," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 74, no. 2, 2012, pp. 337–360.

[26] J. Cadena, G. Korkmaz, C. J. Kuhlman, A. Marathe, N. Ramakrishnan, and A. Vullikanti, "Forecasting social unrest using activity cascades," PloS one, vol. 10, no. 6, 2015, p. e0128879.

[27] L. Zhao, F. Chen, J. Dai, T. Hua, C.-T. Lu, and N. Ramakrishnan, "Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling," PloS one, vol. 9, no. 10, 2014, p. e110206.

[28] G. Korkmaz, J. Cadena, C. J. Kuhlman, A. Marathe, A. Vullikanti, and N. Ramakrishnan, "Combining heterogeneous data sources for civil unrest forecasting," in Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on. IEEE, 2015, pp. 258–265.

[29] N. Ramakrishnan et al., "'Beating the news' with EMBERS: forecasting civil unrest using open source indicators," in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014, pp. 1799–1808.

[30] "Hungarian Algorithm," https://en.wikipedia.org/wiki/Hungarian_algorithm, [Online; accessed 01 July, 2017].

# Approach for Identification of Artificially Generated Texts

Katerina Korenblat and Zeev Volkovich

Department of Software Engineering
ORT Braude College
Karmiel, 21982, Israel
Email: katerina@braude.ac.il, vlvolkov@braude.ac.il

***Abstract*—The paper is devoted to a new method for the identification of the artificially composed scientific papers. We consider this problem from the general point of view of the writing style. It is natural to suppose that the style of artificial generated manuscripts has to be substantially different from this one of the human generated articles because the human writing process is established in inherently another manner. The Mean Dependency Distance introduced in previous authors' works is used to quantify the writing process developing. A set of artificially generated manuscripts is taken and the distance values are calculated to sequential chunks of all papers. A suspected document is also divided into chunks, and a version of the known $KNN$ method is applied together with a distance-based outlier detection method to classify it as a real or a fake document. The provided numerical experiments demonstrate high ability of the method to distinguish between two types of documents.**

***Keywords–Scientific Frauds; SCIgen; Classification.***

## I. INTRODUCTION

In 2005, three computer science Ph.D. students at the Massachusetts Institute of Technology-Jeremy Stribling, Max Krohn, and Dan Aguayo proposed a program, named SCIgen, intended to produce senseless manuscripts in the computer science field. Afterward, a group of scientific document generators was invented, including SCIgen-Physic concentrating on physics, Mathgen focusing on math and the Automatic SBIR (Small Business Innovation Research) Proposal Generator dealing with grant proposal fabrication. Initially, the generators were created as hoaxes with the aim to unmask scientific conferences that really rip off researchers with publication and fees. At the first glance, generated papers appear to be sensible, because they are structured to have all needed components of a paper, such as an abstract, an introduction, graphs, diagrams, citations and so on. The papers are reasonably organized employing context-free grammar and could confuse inexperienced persons. Not formally speaking, the named generators have learned the overall rules commonly used during during writing scientific papers and successively imitate this process. Each person can compose a fake paper using the site [1].

Some articles studied automatic identifying of SCIgen papers. For example, the problem was respected in [2] by proving of the external references. A paper is considered as artificial if a portion of its unrevealed references is sufficiently large. The paper [3] deals with distribution of a papers keyword inside the document, which is natural expected to be appropriately uniform. In [4], a compression profiles of texts are analyzed. The conclusion is based on difference in the compression rate between the authentic and computer generated texts. The ROUGE metrics [5] was used in [6]. Paper [7] along the lines of [8] suggested to use the structural distance between texts. In [9] topological properties of the natural and the generated texts were compared. Different measures to uncover artificial scientific papers were evaluated in [10] and [11].

In this paper, we consider artificial generated manuscripts from the point of view of their own writing style. It is natural to suppose that the style of artificial generated manuscripts has to be substantially different from this one of the human generated articles because the human writing process is completely established in inherently another manner. One of the common viewpoints on the human writing process (see, for example [12]) considers this process as composed of four key elements: planning, drafting, editing, and writing the final draft. Thus, it is natural to presume that dependency between sequential written text parts has remained at the almost uniform level if the text is composed by the same author. On the other site, SCIgen operates with a context-free grammar to produce a text. Essentially, the generator does not compose a paper, but goes along the predefined pattern by randomizing out prior components. An approach quantifying writing style development was introduced in [13]. In this paper, we use a distance between writing styles in order to evaluate dissimilarities between fake and real documents. A set of artificially generated manuscripts is taken, and the distance values are calculated to sequential chunks of all papers. A suspected document is also divided into chunks, and a variant of the known $KNN$ method is applied together with a distance-based outlier detection method to check if it is a real or a fake document.

The remainder of the paper is organized as follows. In Section II, we provide the background on the theory proposed in [13]. The suggested methodology is explained in Section III. Section IV is devoted to numerical experiments. We conclude our paper in Section V.

## II. MEAN DEPENDENCY

Let us consider $\mathbf{D}$ as a collection of texts and get a semi-distance function $Dis$ defined on $\mathbf{D}\times\mathbf{D}$:

- $Dis(\mathcal{D}_1, \mathcal{D}_2) \geq 0$ for all $\mathcal{D}_1, \mathcal{D}_2 \in \mathbf{D}$.
- $Dis(\mathcal{D}, \mathcal{D}) \geq 0$ for all $\mathcal{D} \in \mathbf{D}$.

 ISBN: 978-1-61208-578-4 

It is not suggested that $Dis(\mathcal{D}_1, \mathcal{D}_2) = 0$ implies that $\mathcal{D}_1 = \mathcal{D}_2$. In the framework of our model, we set a chunk size $L$ and consider a document $\mathcal{D} \in \mathbf{D}$ as a series of sequential sub-documents: $\mathcal{D} = \langle \widehat{\mathcal{D}}_1, ..., \widehat{\mathcal{D}}_m \rangle$ of the length $L$. In the formal language theory terminology, $\mathcal{D}$ is the concatenation of $\widehat{\mathcal{D}}_1, ..., \widehat{\mathcal{D}}_m$.

Our perception suggests that a document $\mathcal{D}$ is considered as an outcome provided by "a random number generator" reflecting the writing style of the authors. Aiming to quantify the evolution of a text within the writing process, we introduce the Mean Dependency characterizing the mean relationship between a chunk $\widehat{\mathcal{D}}_i$, $i = T+1, ...m$ and the set of its $T$ "precursors":

$$ZV^{(L)}_{T,Dis}(\widehat{\mathcal{D}}_i, \Delta_i) = \frac{1}{T} \sum_{\widehat{\mathcal{D}} \in \Delta_i} Dis(\widehat{\mathcal{D}}_i, \widehat{\mathcal{D}}), \qquad (1)$$

where $\Delta_i = \left\{ \widehat{\mathcal{D}}_{i-j}, \ j = 1, ..., T \right\}$ is the set of $T$ "precursors" of $\widehat{\mathcal{D}}_i$. To distinguish styles a function measuring dissimilarity among texts pieces is proposed by the following way:

$$DZV^{(T)}_L(\widehat{\mathcal{D}}_i, \widehat{\mathcal{D}}_j) = \qquad (2)$$
$$= \left| \begin{array}{c} ZV^{(L)}_{T,Dis}(\widehat{\mathcal{D}}_i, \Delta_i) + ZV^{(L)}_{T,Dis}(\widehat{\mathcal{D}}_j, \Delta_j) - \\ -ZV^{(L)}_{T,Dis}(\widehat{\mathcal{D}}_i, \Delta_j) - ZV^{(L)}_{T,Dis}(\widehat{\mathcal{D}}_j, \Delta_i) \end{array} \right| .$$

It is easy to see that $DZV^{(T)}_L$ is also a semi-metric. Once $DZV^{(L)}_T(\widehat{\mathcal{D}}_i, \widehat{\mathcal{D}}_j) = 0$ the sub-documents $\widehat{\mathcal{D}}_i$ and $\widehat{\mathcal{D}}_j$ exhibit close relationships with the own previous neighbors and the previous neighbors of another one. From the writing style standpoint the sub-documents appear to be very similar.

Distance function choice is essential in the proposed approach. A relevant distance function may be extracted to reflect writing style attributes. In the text mining domain, it is more acceptable to convert texts into a probability distribution and afterwards to use a distance between them. We suggest that there is a transformation $\mathcal{F}$, which maps the documents belonging to $\mathbf{D}$ into the set $\mathbf{P}$ of the probability distributions on $[0, 1, 2, ...]$, and

$$Dis(\mathcal{D}_1, \mathcal{D}_2) = dis\left(\mathcal{F}\left(\mathcal{D}_1\right), \mathcal{F}\left(\mathcal{D}_2\right)\right),$$

where $dis$ is a distance function (a simple probability distance) defined on $\mathbf{P}$. In the current paper we use the following Spearman's correlation distance function:

$$Dis(\mathcal{D}_1, \mathcal{D}_2) = S(\mathcal{D}_1, \mathcal{D}_2) = 1 - \rho(\mathcal{F}(\mathcal{D}_1), \mathcal{F}(\mathcal{D}_2)),$$

where $\rho$ is the Spearman's $\rho$ (see, [14]), which is calculated for distributions of $\mathcal{F}(\mathcal{D}_1)$ and $\mathcal{F}(\mathcal{D}_2)$ treated as a kind of ordinal data such that the frequency values are regarded as the rank positioning.

As usual, a transformation $\mathcal{F}$ is constructed by means of the common Vector Space Model. This model disregards grammar and the order of terms, but keeps the collection of terms. Each document is described via a terms frequency table in contradiction of the vocabulary containing all the words (or "terms") in all documents in the corpus. The tables are considered as vectors in a linear space having a dimensionality equal to the vocabulary size.

In the Bag of Words Model a document is represented as the distribution of its words. To reduce the space dimensionality, the stop-words are commonly removed. The Keywords Model is an offshoot of the previously discussed model, where a document is represented not as a bag of all terms in the corpus but as a bag of selected words. In the $N$-grams Model the vocabulary consists of all $N$-grams in the corpus. An $N$-gram is a contiguous $N$-character slice of a longer text constructed frequently by means of the symbols occurring in a slide window of length $N$.

## III. Methodology

We handle the considered task in the framework of the one-class classification methodology. One-class classification is based on the presumption that merely data of one of the groups, named the target class, are accessible, although there are no information of the other class (also called the outer class).

In our model the target class is composed from artificially generated papers, while the outliers class is suggested to contain the human written papers. By this way we take a collection of artificially generated papers $\mathbf{D}_0$ and chose the $N$-grams order $N$, the delay parameter $T$ and size of the chunks $L$. All documents from $\mathbf{D}_0$ are divided into chunks having size $L$ and a "cloud" of all chunks $CH(\mathbf{D}_0)$ is constructed as $\left\{ DZV^{(T)}_L(\widehat{\mathcal{D}}_i, \widehat{\mathcal{D}}_j) \right\}$ calculated for all possible, having at least $T$ "precursors", chunks of the documents from $\mathbf{D}_0$. An example of the principal-component analysis plots of such a distance matrix is given in Figure 1, where a cloud is marked in red. A tested document's chunks are marked in blue.

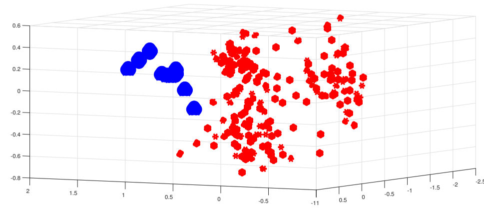

Figure 1. Principal-component analysis plots of a distance matrix

As can we see, the points corresponding to a test document are actually outliers in the red marked documents of the artificially generated papers, because they are located outside of their cloud. There are a lot of methods of the one-class classification (see, for example [15]). We use in this paper two of the most common approaches:

1) The $KNN$ classification, where a text segment is assigned or not to the class by a majority vote of its $k$ nearest neighbors. This algorithm is very intuitive one natural suggesting that each segment is similar to its nearest neighbors.
2) Distance-based outliers ($DBO$) detection:
    a) Determine a central point of $\mathbf{D}_0$ :

$$\widehat{\mathcal{D}}_0 = \underset{\widehat{\mathcal{D}} \in CH(\mathbf{D}_0)}{\arg\min}\ mean\left( DZV^{(T)}_L(\widehat{\mathcal{D}}_0, \widehat{\mathcal{D}}) \right).$$

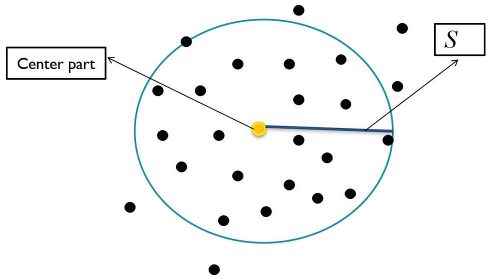


Figure 2. Distance-based outliers detection

b) Calculate $M = mean\left(DZV_L^{(T)}(\widehat{\mathcal{D}}_0, \widehat{\mathcal{D}})\right)$ and $S = std\left(DZV_L^{(T)}(\widehat{\mathcal{D}}_0, \widehat{\mathcal{D}})\right)$, where $\widehat{\mathcal{D}} \in CH(\mathbf{D}_0)$ and $std()$ is a standard deviation function.

c) A chunk $\widehat{\mathcal{D}}$ is recognized as an outlier if $DZV_L^{(T)}(\widehat{\mathcal{D}}_0, \widehat{\mathcal{D}}) > M + S$.

d) A paper is assigned to outliers (a human written paper) if a majority of its own chunks are recognized as outliers.

An illustration of the $(DBO)$ method is given in the scheme presented in Figure 2. Here, the cloud is modelled as a circle having a radius equal to the cloud standard deviation $S$ and the center located at the suggested center of gravity of the cloud. The outliers are associated with this case with point found outside of the circle. Aiming to evaluate significance of a voting in the applied procedures, a $p$-value is calculated within the null hypothesis according to the theoretical fraction $pr$ of the majority voting chunks

$$H_0 : pr = \frac{1}{2}$$

against an alternative hypotheses

$$H_0 : pr > \frac{1}{2}.$$

For the $KNN$ classification $p$-value is found as

$$p = CBIN\left(\widehat{pr} * k, k, \frac{1}{2}\right),$$

where $k$ is the number of the nearest neighbors, $\widehat{pr}$ is a fraction of the majority voting ones, and $CBIN$ is the Cumulative Binomial Distribution.

In case of $DBO$, $\widehat{pr}$ is the observed proportion of the the majority voting chunks within the total number of a document's chunks $m$. Here

$$p = \Phi\left(\frac{\widehat{pr} - \frac{1}{2}}{\frac{1}{2}}\sqrt{m}\right),$$

where $\Phi$ is the cumulative function of the Standard Normal Distribution.

A value of $p$ greater than some predefined threshold, typically $0.95$, indicates a significant voting.

## IV. Numerical experiments

As fake documents, one hundred papers are generated by the SCIGen procedure. One hundred real manuscripts are recovered from the "arXiv" repository [16]. This number of the artificial and human written papers appear to be very reasonable and provides fair results. However, we are going in the future to study a possibility to reduce the size of the training set.

In the texts involved in the experiments any uppercase characters are converted to the corresponding lowercase characters, and all other characters are unchanged. The experiments are provided through the chunk size $L = 100, 200$ and $400$ with $T = 10$. The number $k$ is 10 in the $KNN$-classification.

Fifty artificial papers are used as a training corpus: each document is divided into chunks of size $L$ and the distance values are calculated according to (2). In the first series of the experiments, the real papers were compared with training set using two mentioned approaches. The results obtained from considering real papers are presented in Table I by means of the positive predictive values (precision) calculated within the experiments. Recall that precision is an attained proportion of the true positive results. Almost all $p-$values are properly close to one. It is easy to see that the $KNN$ method exhibits very stable behavior, which is practically independent of the choice of $L$. The achieved outcomes are very precise in all cases. So, almost in all experiments the considered real papers are recognized as human written.

TABLE I. POSITIVE PREDICTIVE VALUES CALCULATED FOR 100 REAL PAPERS.

| *L* | *100* | *200* | *400* |
|---|---|---|---|
| $KNN$ | 0.98 | 0.99 | 0.98 |
| $DBO$ | 0.63 | 0.81 | 0.88 |

In the second experiments series, fifty additional artificial papers were checked against the fifty training ones.

TABLE II. POSITIVE PREDICTIVE VALUES CALCULATED FOR 50 ADDITIONAL ARTIFICIAL PAPERS.

| *L* | *100* | *200* | *400* |
|---|---|---|---|
| $KNN$ | 1 | 1 | 1 |
| $DBO$ | 1 | 1 | 1 |

So, each fake document is acknowledged with probability one.

## V. Conclusion

The article proposes a new technique for recognition artificially generated scientific papers resting upon written style characteristics. Texts are split into chunks and shown by means of a histogram shape defined via the 3-gram frequency ranks. Then, the Mean Dependency describing the mean rank correlation of a sub-document with its numerous precursors provides a distance amid chunks styles. Two classifiers by means of this distance within the $KNN$ and the distance-based outlier detection methodologies are constructed and tested. It turns out that a $KNN$ based approach trained on a sufficient number of fake papers is capable almost surely to distinguish between fake and real papers. The second classifier demonstrates less robust results. We are planning to extend our

 ISBN: 978-1-61208-578-4 

method aiming to construct new classification rules using one-class and two-class classification approaches.

## REFERENCES


[1] "Some Webpage," URL: https://pdos.csail.mit.edu/archive/scigen/ [accessed: 2017-07-02].

[2] J. Xiong and T. Huang, "An effective method to identify machine automatically generated paper," in Knowledge Engineering and Software Engineering, 2009. KESE'09. Pacific-Asia Conference on. IEEE, 2009, pp. 101–102.

[3] A. Lavoie and M. Krishnamoorthy, "Algorithmic Detection of Computer Generated Text," eprint arXiv:1008.0706,2010arXiv1008.0706L, AUG 2010, aRXIV.

[4] M. M. Dalkilic, W. T. Clark, J. C. Costello, and P. Radivojac, "Using compression to identify classes of inauthentic texts," in Proceedings of the 2006 SIAM Conference on Data Mining, 2006.

[5] C.-Y. Lin and E. Hovy, "Automatic evaluation of summaries using $n$-gram co-occurrence statistics," in Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1, ser. NAACL '03. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003, pp. 71–78. [Online]. Available: http://dx.doi.org/10.3115/1073445.1073465

[6] D. P. F. Labbé, Cyriland Labbé, Detection of Computer-Generated Papers in Scientific Literature. Cham: Springer International Publishing, 2016, pp. 123–141.

[7] U. Fahrenberg, F. Biondi, K. C. J. Kongshøj, and L. Axel, "Measuring global similarity between texts," in Statistical Language and Speech Processing Second International Conference, SLSP 2014, Grenoble, France, October 14-16, 2014, Proceedings, 2014, p. 220232.

[8] C. Labbé and D. Labbé, "Duplicate and fake publications in the scientific literature: how many SCIgen papers in computer science?" Scientometrics, vol. 94, no. 1, 2013, pp. 379–396.

[9] D. R. Amancio, "Comparing the topological properties of real and artificially generated scientific manuscripts," Scientometrics, vol. 105, no. 3, 2015, pp. 1763–1779. [Online]. Available: http://dx.doi.org/10.1007/s11192-015-1637-z

[10] K. Williams and C. L. Giles, "On the Use of Similarity Search to Detect Fake Scientific Papers". Similarity Search and Applications, OCT 2015, pp. 332–338.

[11] M.-T. Nguyen and C. Labb, "Engineering a tool to detect automatically generated papers," in BIR@ECIR, ser. CEUR Workshop Proceedings, P. Mayr, I. Frommholz, and G. Cabanac, Eds., vol. 1567. CEUR-WS.org, 2016, pp. 54–62.

[12] J. Harmer, Ed., How to teach writing. Delhi, India: Pearson Education, 2006.

[13] Z. Volkovich, O. Granichin, O. Redkin, and O. Bernikova, "Modeling and visualization of media in arabic," Informetrics, vol. 10, no. 2, 2016, pp. 439–453.

[14] M. G. Kendall and J. D. Gibbons, Rank correlation methods. London: Edward Arnold, 1990.

[15] S. S. Khan and M. G. Madden, "One-class classification: Taxonomy of study and review of techniques," CoRR, vol. abs/1312.0049, 2013. [Online]. Available: http://arxiv.org/abs/1312.0049

[16] "Some Webpage," URL: www.arXiv.org/archive/cs [accessed: 2017-07-02].


 **ISBN: 978-1-61208-578-4** 

# Examining the Interaction Between Fourth Estate and Twitter: An Exploratory Case Study

Balbir S. Barn*, Ravinder Barn†, Franco Raimondi*, and Utsa Mukherjee†
*Dept. of Computer Science
Middlesex University, London, UK
Email: b.barn,f.raimondi@mdx.ac.uk
†Royal Holloway University of London
Egham, UK
Email: r.barn@rhul.ac.uk and
utsa.mukherjee.2015@live.rhul.ac.uk

***Abstract*—The pervasiveness of social media has resulted in increased public involvement in key discussions about social issues, as well as creating greater affordances for individual expression and collective mobilisation. In December 2012, the rape and murder of a 23-year-old Indian student in New Delhi, India, was followed by widespread condemnation and public action organised and coordinated through social media (Barn 2013). In March 2015, a controversial BBC documentary, "India's Daughter", about the incident was broadcast despite restrictions imposed by the Indian Government. This paper explores the interplay between mainstream media (the so-called Fourth Estate) and Twitter through a case study analysis using computational techniques to analyse 250000 tweets collated following the broadcast of the documentary. In particular, we apply the theory of postcolonialism to understand the dynamics of this interaction. Issues around implications for conducting inter-disciplinary social media research are also discussed.**

***Keywords–Postcolonialism; India's Daughter; Fourth Estate; Social Network Analysis; Social Media; Twitter***

## I. Introduction

As the most populous democracy in the modern world, India has witnessed an increasing growth in the use of the internet in general, and social media in particular. Although accurate statistics are difficult to obtain, estimates of the micro-blogging site of 140 characters, Twitter, range from 23 million to 35 million [1]. This figure has more than doubled in the last 3 years. According to a collective called 'India on the Internet 2014', Twitter users in India total 35 million while 125 million people are on the social networking site (SNS), Facebook. Further, it is estimated that almost 9 out of 10 web users in India visit a social networking site. In a climate of smart phones and their applications, and SNSs such a figure is not so surprising. A core features of such social media sites is their reliance on individual users for content creation and active user involvement.

Twitter is notable in that it has rapidly become important and popular as key tool for organising and generating communication for protestors around the world [2].

In addition to the organising and communicating aspects of Twitter, researchers have also commented on how Twitter and social media in general is also being used to reconstruct and extend journalism and notions of what constitutes a Habermasian public sphere [3], [4]. That is, the realm of social life in which something approaching public opinion can be formed and where access is guaranteed to all citizens. The outcome is that networked individuals have the capacity to use social media to enhance their role in news production and dissemination to achieve a growing independence from the Fourth Estate (mainstream media) [5].

In South Delhi, India, the 16 December 2012, rape and murder of a 23 year old physiotherapy student by six men, marked a watershed moment where some commentators asked whether this heralded an 'Indian Spring' [6]. The street protests across the nation in which social media was said to have played a part were described by some as the new 'unifying force' [7] through the formulation of a shared public opinion on social media.

Several years after the Delhi rape, a BBC documentary titled *"India's Daughter"* was broadcast on 4 March 2015 in the UK and on 8th March 2015 in New York. The broadcast of the documentary directed by Leslee Udwin was controversial, in that, the Indian Government sought to have it banned and the BBC chose to bring forward the broadcast to an earlier programme slot. Early indicators of the controversial aspects of the documentary were immediately brought to the foreground. These included: the extent to which mainstream media occupied the so-called egalitarian and democratising space of social media; the postcolonial texture of the debate and the overall sense of how western media handled the case under question. Twitter naturally formed predominant backdrop to this broadcast given its initial role in mobilisation of public opinion in the original 2012 event [6], [7], [8].

Within a context where the mainstream media and social activists now largely occupy this micro-blogging space [7], research that examines the interplay between mainstream media (Fourth Estate) and social media (so called Fifth Estate) through specific case study instances can have public policy implications. Hence, the research reported in this paper makes a key analytical contribution of public reaction, through tweets, to the broadcast of the documentary on the BBC in the UK and on Youtube, Vimeo and other sources in India.

The remainder of the paper is structured as follows: In Section 2, we outline key concepts and debate around postcolonial theory given that postcolonialism was an important emergent theme. Crucially, the historical relationship between the two countries where the incident occurred and the film was shown provided a particular salience and backdrop to the analysis of this paper. In Section 3, we present the aims and details of the research methodology we have used. Section

 ISBN: 978-1-61208-578-4 

4 presents overview results arising from the blog analysis and computational analysis of the collected tweets. Section 5 entails a discussion of some tweets followed by commentary on the validity of the results. Finally in Section 6, we present concluding remarks and outline some further research considerations.

## II. Related work and underpinning theory

As noted earlier, Twitter is both important and popular as key tool for organising and generating communication for protestors around the world [2]. Examples of where Twitter has played a significant role include: the Iranian protests of 2009-2010 [9], the so-called Egyptian revolution of 2011 [10] and also the various Occupy protests that took place around the world [11]. It is also clear that that the messaging technology is viewed differently depending upon temporality and context. Hence, it is seen as subversive by autocratic regimes, as well as a suitable technology for surveillance [12]. In an analysis of 104,059 tweets related to the Delhi rape incident and social protests that took place across urban India, in line with previous scholarship, Ahmed and Jaidka (2013) conclude that traditional media still plays a pivotal role in disseminating information [7]. For instance, the authors report that less than 10% of the tweets were actually from ordinary citizens / individuals. Such a finding certainly lends credence to previous observations that have questioned the egalitarian, and democratising promises of such space [13]. Questions also arise as to whether a new public sphere is being reconstructed where ordinary citizens really do have an opportunity to form public opinion. Hence, it makes sense to also provide an understanding of western media's handling of the case under question. This is particularly important to help ground the response of the postcolonial society and to also more properly explore the role of the impact of social media on public policy.

Over the last several decades, postcolonial theory has emerged as a major intellectual critical approach. The theory is generally regarded as having been founded on the contributions of key writers including Frantz Fanon, Edward Said, Gayatri Spivak, and Homi Bhabha. Primarily, postcolonial theory seeks to problematise key historical and contemporary notions, structures and processes including colonialism, race, ethnicity, culture, racism, gender, identity, inequality, and globalisation. In short, the theory seeks to 'critique and aims to transcend the structures supportive of Western colonialism and its legacie' [14]. The watershed moment for the polemics of postcolonialism was the publication of Edward Said's *Orientalism* in 1978 [15]. In *Orientalism*, Said meticulously brought out, through close textual studies, the prejudices about and biases against the non-West that informed the colonial discourse and its meaning productions. He showed how the non-West or the orient came to occupy the space of an exotic 'Other' in the canon of Western knowledge and how this 'orientalism' as a discourse was responsible in justifying the colonial and imperial projects of the Western powers. Postcolonialism, therefore, became the rallying point to challenge the presumptions of Western knowledge systems, to comprehend the epistemologies of the non-West, to create a space where, as Gayatri Chakravorty Spivak puts it “the marginal can speak and be spoken, even spoken for” [16]. Such an epistemological framework is indeed one of the key components of postcolonial theory. Others include its critique of power in the forms of economic/cultural/economic/political/ideological domination (both historically and in the present time), its stance on the processes of otherisation, and essentialism. In the context of postcolonial theory, otherisation refers to a process by which one group uses social and psychological means to exclude or marginalise another group by focussing on differences. Whereas essentialism is understood to be the essence or "whatness" of something. In postcolonialism, essentialism implies the action of how a colonising power decides what is and isn't a particular identity. More often than not, differences and/or commonalities between groups may be overlooked to maintain a power relation.

Given this backdrop, in an examination of verbal and visual texts in United States mainstream news media reporting of the Delhi rape case, Durham argues that India / Third World is 'represented as a primitive and undisciplined space populated by savage males and subordinate women' [17]. She further asserts that in the geopolitics of sexual assault, the USA news media reinscribed social geographies of power and sex in terms of gender. Such an ethnocentric framework portrays the Third World woman as oppressed and lacking in agency, and the nation-state as incompetent and complicit in her subordination. The mediated deployment of space and place and Delhi in particular as the 'rape capital' of India serve as a key signifier of the political economy of gender and sexuality, and hence the process of ranking one society over the other.

In her analysis of over 1500 USA mainstream news articles published over a period of two months, following the December 2012 rape incident, Roychowdhury (2013) argues that through its coverage, the news reporting not only created a polarity between the new and old India within a new-liberal consumer world; but also stressed the 'notions of Western gender progressivism' as evidenced through its language including words such as 'traditional societies', 'medieval', 'rape as a weapon of power against modernity'. Here, in spite of the evidence on crimes of rape against women in the west, western space with its so-called modern cosmopolitanism is presented as safer for women. The December 2012 case is used as a platform to present a dichotomy of the modern Indian woman victim, and the backward / savage /misogynist brown man. Roychowdhury cites Spivak‘s 1988 writings, and argues for its ongoing appeal as witnessed in western media, that is, "white men saving brown women from brown men." [18].

We use this texturised context of postcolonialism to examine the extent to which Spivak‘s theoretical framework can be employed to explore the interaction between the Fourth Estate and Twitter in the context of the BBC documentary, India's Daughter. Indeed, the broadcast of this Fourth Estate film, and its discussion on Twitter provides a useful anchor to extend the postcolonial lens referred to by scholars, such as Durham and Roychowdhury.

## III. Aims and Methods

In this study, our primary aim was focused on the interaction between the Fourth Estate and Twitter. In doing this, we sought to apply the theory of postcolonialism to understand the dynamics of this interaction. Additional areas of interest included an identification and exploration of the debates and discussion generated as a consequence of this controversial BBC documentary. To this end, we employed a mixed-methods approach to help understand the situation namely: a series of

 ISBN: 978-1-61208-578-4 

blogs written at the time of the broadcasting of the film in the UK, and USA in March 2015; and then the collection and analysis of tweets over a period of 4 weeks (3 March 2015 - 3 April 2015). Notably, the film was broadcast during this period to coincide with International Women’s Week. The topicality and contemporaneous nature of the study required drawing upon those social media blogs, written within a week or so of the broadcasting of the film. The blogs were read and analysed manually by two of the authors and led to the identification of the dominant themes used in the subsequent analysis. These blogs constituted not only as part of our data collection, but they were also useful in contextualising and in making sense of our Twitter data. Sixteen blogs from prominent bloggers were analysed. Most of the bloggers were female (11), and only three were male. Two of the bloggers were unknown.

A key challenge in conducting social media based research is the lack of standard approaches in appropriate methods for data collection and analysis. This concern tends to be further compounded by a limited range of integrated tools to support research methods that can enable the full range of types of analyses required. The Collaborative Online Social Media Observatory (COSMOS) is an example of a distributed digital social research platform that addresses these requirements [19]. However, at the time of this research, the tool was not readily available and furthermore did not integrate with our efforts at developing a learning set through manual analysis. Other tools such as Prometheus is a peer-to-peer service that collects social data from a number of sources and applies social inferencing techniques, but it is mostly concerned with privacy-aware social data management [20].

Given these concerns, we chose to access the Twitter data stream using the published Twitter Application Programming Interface (API) via our own bespoke software. Twitter offers a streaming API that can be filtered on keywords; in our case we employed the following list: “IndiasDaughter”, “Leslee Udwin”, “Udwin” and “banbbc”. The script was kept running to collect tweets that included the keywords from 3rd March 2015 to 3rd April 2015 following the broadcast of the documentary. Over 254,000 tweets were collected amounting to around 1GB of data. Such a volume of data requires computational approaches to analysis. Figure 1 provides an illustration of the general steps in our method.

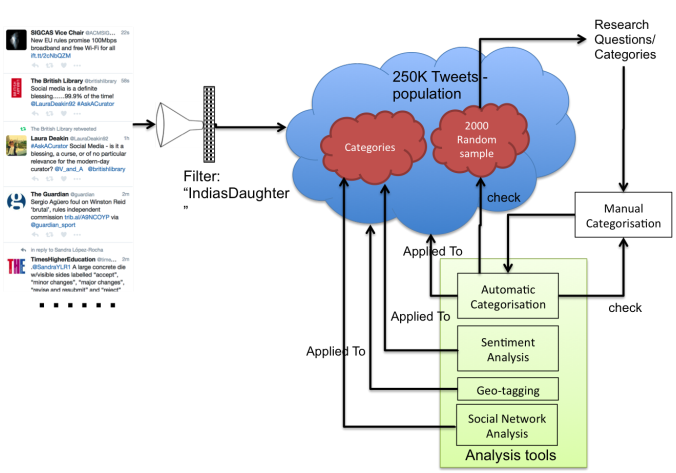


Figure 1. Method Overview.

Small scale analysis of tweets of say less than 10000 is relatively straightforwardly done by human processing. Computational approaches provide additional insights that would not necessarily be possible by manual analysis. Given the volume of the tweet set, we were interested in several types of analysis. These analytical tools included: Automatic categorisation: the use of machine learning to categorise text or other data; Sentiment analysis: the use of natural language processing and text analysis to identify attitudes of a respondent with respect to a topic; Geo-tagging: plotting the location information of tweets on a map; and Social network analysis: the use of network and graph theory to investigate social structures. We discuss the use of the analytical tools below using the diagram in Figure 1 to provide an overall context.

## IV. Results

As indicated earlier, our research questions have centred around several exploratory areas. The blogs, existing literature and our own research questions prompted and influenced these exploratory areas which would become coding categories in our thematic analysis of the tweets.

### A. *Analysis of the Blogs: Emergent themes/categories*

In the analysis of blogs from March 2015, several divergences were identified. Firstly, some bloggers set out to support the ban and to justify it [21]. Others proposed taking a legal stand on the matter and argued that as the case is still subjudice, the telecast should be postponed until such time as the judgement is pronounced by the court, but in no way supporting the ban [22]. Bloggers also chose to challenge the ban, ask for it to be lifted immediately, and the telecast to take place as per the schedule [23], [24]. Others took an informed and critical stand, and commentated from various perspectives such as feminism, postcolonialism or even with India’s general use of "bans". Arguments for condemning and supporting the ban in the same breadth were presented [25], [23]. An analysis of the blogs helped generate a useful framework that could be applied to our Twitter dataset. A total of 7 prominent themes were identified. These included notions of legality of broadcasting the documentary (Legality/Ban), the postcolonial mindset of the film and the response from others (Postcolonialism), representations and discussions about the lawyers involved (Lawyers), contemporary feminist thinking in India (Feminism), the role of traditional media in discussions, in this case BBC and NDTV (mainstream media), representations and discussions about the role of Government of India (Government), and finally the role and value of punishment (Punishment).

### B. *Categorisation of tweets*

Thematic analysis or categorisation is a powerful qualitative data analysis tool. The challenge is to deploy it for 1 GB of data. We elected to use machine learning techniques and the use of training sets. A random sample of 2000 tweets were extracted from the tweet population and classified manually against the categories listed above by two of the researchers. Where there was discrepancy, discussion was used to agree a final classification. This tweet set of 2000 tweets was used as a "Training Set" to refine / parameterise machine learning algorithms which were then used on the entire tweet population to categorise the 254K tweets. We partitioned the 2000 training set using 3 folds. We used NTLK, a Python based toolkit for natural language processing [26]. This software comes with an open source library and toolkit for natural language processing to do stemming and tokenisation, using all the words as features. We have employed a Naive Bayesian classifier to build our model. The table in Figure 2 overleaf summarises the categorisation results. This training set had around 72% accuracy (manual versus automatic categorisation). This is consistent with other research [27]. We finally applied the model to the whole dataset of 254K tweets. Tweets related to postcolonialism amounted to 26,816, representing 10.5% of the overall total and the second largest of the analytical categories. The largest category centred around tweets about legality and banning of the broadcast. More importantly, the postcolonialism tweets amounted to 23% of the tweets that were classified against the desired classifications by the machine learning algorithm. We used the 'Other' category to denote discussion that did not fall into the categories of interest.

### C. Sentiment Analysis

The use of natural language processing and text analysis to identify attitudes of a respondent with respect to a topic is a popular analytical tool used in tweet analysis. Both the original incident (through the brutality of the crime) and the subsequent controversial aspects of the televised documentary generated a wide range of emotions and efforts to assess the overall sentiment was deemed appropriate. We used the open source vaderSentiment0.5 tool [28] to conduct a sentiment analysis of the tweet data set. VaderSentiment represents sentiments on a scale from -1 to 1 representing negative sentiment at one end (-1) and positive sentiment at the other end (1). When the full set of tweets were subjected to a sentiment analysis using VaderSentiment, we found that tweets related to postcolonialism were ranked 3rd in association with negative sentiment. The overall compound sentiments for all categorisations is shown in Figure 2.

| Analytical Category | (n) |
|---|---|
| Legality/Ban | 42,870 |
| Postcolonialism | 26,816 |
| Lawyers | 17,150 |
| Feminism | 12,129 |
| Media | 8,757 |
| Government | 4671 |
| Punishment | 2186 |
| Other | 139, 841 |
| Total | 254,430 |

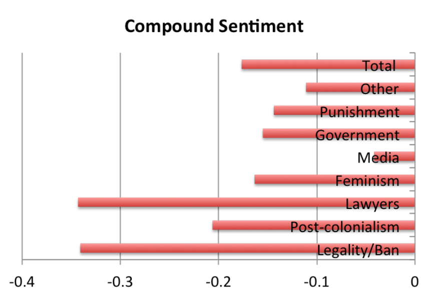


Figure 2. Automatic categorisation and overall sentiments.

### D. Social Network Analysis

In this paper we are predominantly concerned with the postcolonial texture of the debate surrounding the broadcasting of the film. Hence, social network analysis on the tweet set associated with postcolonialism was conducted using two open source social network analysis tools: NodeXL [29] and Gephi [30]. The Postcolonialism tweet set (JSON file) was transformed using bespoke scripts into a form readable by NodeXL. In this data set, nodes are Twitter users and the edges represent tweets that can either be retweets or mentions. This data set was further cleaned by merging duplicate edges and the addition of weights to reduce the edge count. NodeXL was used primarily as a means for creating the GraphML format for use in the visualisation within the Gephi toolset. Gephi is better supported on MacOS and we were able to compute and visualise various social network analysis descriptive statistics. Within Gephi, the data set comprised 2000 nodes and 13243 undirected edges. Modularity computations were performed on the network. These measured the strength of division of a network into communities. We used the modularity algorithm included in Gephi [31] and produced seven communities of interest. Each of the top four communities (in size) were centred around key Fourth Estate actors such as @BBC, @NDTV @BDUTT, @BBCIndia and @TimesOf India. Also apparent, was how these same 4th estate actors (i.e. traditional media) were also ranked highly in a range of network centrality measures such as Betweenness Centrality and Eigenvalue centrality.

*Degree Centrality* is a measure of a node with respect to its in-bound connections and its outbound connections. If a node/actor receives many ties they are often regarded as prominent or important. Nodes that have a high out-degree are actors that are influential.

*Betweenness Centrality* is a measure of node that is based on the extent to which a node falls on the geodesic paths between other pairs of nodes in the network. In social network analysis, nodes with a high value for betweenness centrality are an indication of influence on information flow in a network. Hence a node with a high value is an important conduit for information flowing between nodes in the network.

*Eigenvalue Centrality* considers in-bound and out-bound connections and also the node's connection to other important nodes. Hence, the measure is seen as an indicator of the power of the node.

Tiryakian et al. note that "Individuals with high betweenness centrality tend to be influential because they are well informed and can affect the flow of information in a network. As a result, they are often information gatekeepers." [32]. For example, @BBC was top-ranked for both Eigenvector (0.00929400), and Betweenness centrality (13556362.612). We also observe that the use of Eigenvalue centrality to denote power is open to debate and recent results have indicated that in Twitter, users with high eigenvector centrality need not be influential users [33]

The film was shown in both USA and the UK of which the latter has a postcolonial relationship. Analysis of geo-tagged tweets was not considered meaningful given the low numbers of geo-tagging. Other techniques such as automated analysis of Twitter handle biographies would have been a potential route for analysing how sentiments and other issues vary between the countries where the films where shown. Our future work will incorporate such approaches.

Our results for the centrality statistics are shown in table 1. These data are the non-normalised results. The top 25 results are shown. From the table it is clear that there is considerable overlap between the centrality measures of various nodes. @BBC for example is the most powerful node in the network and one with the highest betweenness measure. Several BBC based Twitter accounts feature as important conduits for information flow. Several of the nodes were Hindu nationalists,

 ISBN: 978-1-61208-578-4 

TABLE I. CENTRALITY MEASURES

| Vertex | Eigenvector Centrality | Vertex | Betweenness Centrality |
|---|---|---|---|
| BBC | 0.00929400 | BBC | 13556362.612 |
| lesleeudwin | 0.00635000 | BBCIndia | 9210704.414 |
| YesIamSaffron | 0.00344600 | AlanaBowker | 7989346.749 |
| MahaveerM_ | 0.00344600 | lesleeudwin | lesleeudwin |
| Sootradhar | 0.00304900 | LutyensInsider | 4750346.108 |
| ndtv | 0.00301000 | Sootradhar | 3742481.265 |
| TRobinsonNewEra | 0.00290400 | BBCWorld | 2470768.572 |
| dibang | 0.00289000 | gvicks | 2415990.807 |
| rishibagree | 0.00261900 | ndtv | 2385230.780 |
| LutyensInsider | 0.00260200 | vicky895 | 2364758.802 |
| PMOIndia | 0.00250600 | dibang | 2340111.494 |
| DrJwalaG | 0.00228700 | shaliniscribe | 2269173.050 |
| BBCWorld | 0.00228700 | sunnysingh_nw3 | 2145012.461 |
| thekinshu | 0.00193500 | M_Lekhi | 1981712.913 |
| Keisar_ | 0.00184800 | Keisar_ | 1967994.765 |
| mediacrooks | 0.00169900 | rishibagree | 1806544.314 |
| narendrapjoshi | 0.00168700 | MahaveerM_ | 1738846.955 |
| mariawirth1 | 0.00155700 | rvaidya2000 | 1634460.299 |
| rvaidya2000 | 0.00149300 | mariawirth1 | 1634460.299 |
| sankrant | 0.00147700 | seemagoswami | 1523056.837 |
| AskAnshul | 0.00143200 | YesIamSaffron | 1458497.885 |
| BDUTT | 0.00142300 | SuchayVora | 1275945.185 |
| surnell | 0.00138700 | radhabharadwaj | 1259731.335 |
| htTweets | 0.00134600 | PMOIndia | 1206559.332 |
| rohitagarwal86 | 0.00129200 | anniegowen | 1176381.602 |

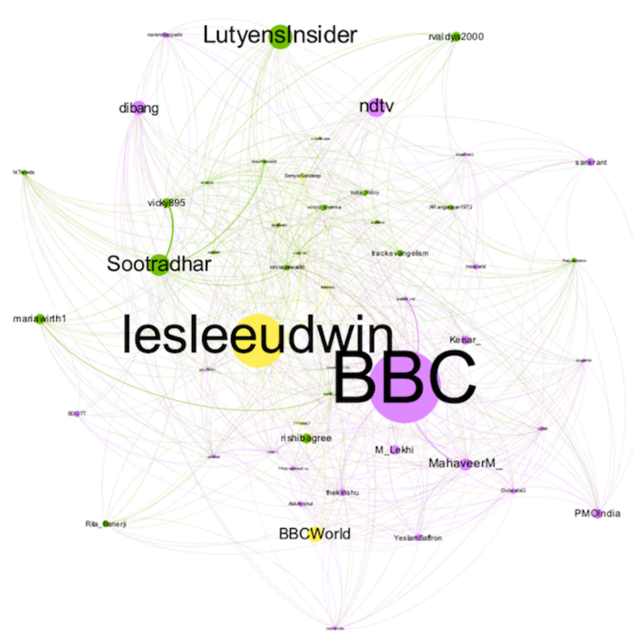


Figure 3. Power Diffusion

Others were bloggers such as @Sootradhar and @thekinshu. Presenters on Indian television programmes were represented and included @dibang, and @BDUTT for example. Most were individuals.

Figure 3 above depicts the core of the 2000 node undirected network by limiting the map to vertices (nodes) that have degree range of 25-89 edges and that also have an Eigenvalue > 0.1111762270211876 (normalised). These parameters were used primarily for presentation purposes. Additionally, the vertices are scaled by Betweenness centrality to indicate the roles that vertices are playing in brokerage and diffusion of information. The various communities to which vertices belong are also indicated by the colour.

## V. Discussion

Our results indicate that mainstream media appear to be effective in utilising the communication spaces created by social media. Thus, for example, the BBC and NDTV Twitter accounts features in both measures in Table 1. However, we can observe some impact of Fifth Estate through the centrality measures observed for some influential bloggers (@Sootradhar and @thekinshu) but noticeably not from those bloggers who were perceived to have set the agenda through their articles. The positioning of social media as a force for dissemination and a viable growing independent "Fifth Estate" as an alternative to mainstream media based journalism [34] should be questioned.

The BBC finds mention in around 50% of the tweets that were posted on the issue of the ban (categorised as Postcolonialism), but there were very few tweets from the various BBC official handles in that data set (<10). The BBC World News, with the handle @BBCWorld, on the other hand retweeted some of the tweets coming from @BBCIndia. These include the following: @BBCIndia: Letter from Director of BBC Television to Indian broadcast ministry on #IndiasDaughter. The letter that was attached to this tweet, was the one written by Danny Cohen, the Director of Television at the BBC, to the Joint secretary at the Ministry of Information and Broadcasting, Government of India. The following quote from the letter helps disambiguate the BBC's position:

*...We appreciate your concern but we feel India's Daughter has a strong public interest in raising awareness of a global problem and the BBC is satisfied with the editorial standards of the film... The purpose of including the interview with the perpetrator was to gain an insight into the mind-set of a rapist with a view to understanding the wider problem of rape and not just in India.* [35]

Further, the letter goes on to make clear whilst there is a notional respect for Indian legal jurisdiction, the BBC would bring forward the broadcast of the documentary to an earlier slot. Reasons are not given, and one possible inference is potential legal action preventing the broadcast.

*...It should be noted, although the BBC is happy to take your views into consideration, we are not planning to transmit the film in any territory which lies under Indian legal jurisdiction...after lengthy and careful consideration we have decided to show India's Daughter on BBC Four in the UK at 22:00 tonight.* [35]

What is attempted in this letter has a direct bearing on our postcolonial understanding, for it redraws the battle-lines for this debate and depicts 'rape', the central issue of the film, as a *global* and not an India specific malaise. And in so doing, it attempts to steer away from any alleged bias in the representational politics despite the obvious conflation in the title of the documentary which asserts a national identity to the malaise of rape. The letter's reference to *jurisdiction* appears to underplay the ubiquity and reach of social media, and seems to embolden the role of mainstream media. The Twitter discussions amplify some of the positions identified by the bloggers. Thus, Mehra notes:

*There are many questions regarding why some sections of the widely-read Western media are so obsessed with reporting and analysing rapes in India while often conveniently ignoring* *similarly grievous nature of the problem in their own backyards.* and asks: *Can you imagine if an Indian film-maker were to make a documentary on the thousands of cases of sexual abuse and 'grooming' of young girls in Rotterham (sic) between 1997-2013, and title it "Britain's Daughter"?* [36].

The blogs and tweets from other central nodes (including bloggers and Hindu nationalists) not only accused BBC of being prejudiced in their framing of the narrative, but also came down upon the broadcaster for doing so at the expense of not giving adequate attention to gruesome incidents and issues at 'home'. For example:

***@YesIamSaffron*** How Many Documentaries ve u Seen @BBC Reporting US Rapes or Biggest Child Exploitation Scandals in UK? #BanBBC #UHF http://t.co/hglD7lfHOP (297 RTs)

***@thekinshu*** Dear @lesleeudwin can you pl tell me why you afraid to make a Documentary on your own Rape Case in London.? #BanBBC

The above tweets demonstrate a strong dynamic of nationalism, and the 'other' that portrays the BBC and the documentary film maker as an oppressive force. There is a challenge to the claim made by the 'other' to be the holder of morality, and superiority. Moreover, the 'other's' stance on human rights and gender is also questioned in highlighting its own social problems. Crucially, the BBC and Leslee Udwin serve to represent the Western 'other'. In doing so, they become emblematic of the UK, the ex-colonial power and its representation of postcolonial India.

A remarkable anomaly was the hijacking of the debate by the UK anti-muslim, right-wing party leader, Tony Robinson and how Hindu nationalist Twitter users used a tweet (retweeted over 2000 times) from Tony Robinson for their own ideological stance (see @YesIamSaffron above) bringing to the foreground a curious postcolonial paradox.

***@TRobinsonNewEra*** We are in the middle of the biggest rape and child exploitation scandal in our countries history, and the @BBC are focusing on india's rapes (2078 RTs)

It would seem that networks and how power is transferred through such networks appear to be structured by colonial and national tensions and complicated by religion, and other social divisions.

### *A. Threats to Validity*

Internal validity issues centre around whether the micro posts (tweets) are appropriately categorised by the machine learning algorithms. Here, care was taken to use a manual process of categorisation of 2000 randomly selected tweets to help develop the training set for the machine learning software. While there is confidence that the categorisation has operated at the reference 72% accuracy [27] it is noted that a significant percentage of tweets could not be automatically categorised. Note, however, no claims on external validity (wider generalisations to different domains) are being made.

Methodological notions of validity, reliability and repeatability present a concern for much of social media research as it is challenging to be definitive that data collected from social media is actually representative of the phenomena of interest. Twitter users commenting on a particular phenomena are self-selecting. Such concerns can be partially mitigated by the use of large scale analysis (using large data sets) but nonetheless risks such as social media being generative of the behaviours it aims to document are paramount [37]. A further concern is related to decisions around treating journalists as individual citizens or as part of the overall machinery of the Fourth Estate. We view the latter as a more representative definition.

Social media research requires inter-disciplinary thinking, but arriving at a common ground whereby a sociological theory can be adequately expressed for computational purposes is an ongoing research challenge [38].

Social network analysis, and in particular the centrality measures can offer some insight about power diffusion in networks, but as noted earlier, are open to debate. Eigenvalue centrality for example, unless correlated with inbound/outbound data may not be a good indicator of power.

Furthermore, there are limited open source seamless software tool chains that addressed the types of analyses that we utilised. There are risks in moving data between software tools. Importantly, we restrict our claims to the data that we have collected and make no generalisations for other contexts.

The ethics of using data published in social media is also of concern. The approach taken in this paper considers two key dimensions, risks of identification / disclosure to users and ethical risks around the content of the micro-blogs. The data collection (both the blogs and the tweets) are from users who would be classified as low risk users as either the user is not identifiable from a Twitter profile or is from a public, official or bot account. Ethical risks related to the content of the tweets are also limited. While the content is at times provocative and antagonistic, the classification of the users as low risk does not warrant opt in permissions before publication. Possibilities of masking identities can address confidentiality concerns, but it is important to note that Twitter Terms and Conditions state that tweets must be given in their original form and attributed to the individual who posted the tweet. Furthermore, informed consent becomes near impossible when dealing with data at large scale. The dominant use of hashtags by Twitter users also supports the notion that such users were broadcasting their thoughts specifically on a subject in a public discussion [39].

## VI. Conclusion

Social media is becoming an important communication channel in the modern world, however its role as a democratic tool, its reach and therefore direct impact remain questionable in terms of policy formation. The recent presidential election in USA has certainly brought to the foreground, the role of social media and the need for policy discussions [40]. Our study shows that the traditional media and their components such as individual journalists continue to play a central role in these new spaces. The influence and reach of the ordinary citizen is less well pronounced so it is uncertain that the so-called Fifth Estate is really coming to the fore. Methodological concerns remain challenging. For example, common grounds whereby a sociological theory, for example, postcoloniality (as in this paper) can be adequately expressed for computational purposes remains an open question. Software tools that can support both sociological reasoning and computational analysis in a linked and coherent way has the potential to have a significant impact on addressing the issues of validity raised. Some of our future research effort is directed at developing a software tool

 ISBN: 978-1-61208-578-4 

chain that social scientists will be able to use independently of computer scientists.

## References


[1] I. Facts, “Topic: Internet usage in India,” 2016, Retrieved: 10-07-2017. [Online]. Available: https://www.statista.com/topics/2157/internet-usage-in-india/

[2] M. Tremayne, “Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street,” Social Movement Studies, vol. 13, no. 1, 2014, pp. 110–126.

[3] J. Habermas, The structural transformation of the public sphere: An inquiry into a category of bourgeois society. MIT press, 1991.

[4] V. Belair-Gagnon, S. Mishra, and C. Agur, “Reconstructing the Indian public sphere: Newswork and social media in the Delhi gang rape case,” Journalism, vol. 15, no. 8, 2014, pp. 1059–1075.

[5] N. Newman, W. H. Dutton, and G. Blank, “Social media in the changing ecology of news: The fourth and fifth estates in Britain,” International Journal of Internet Science, vol. 7, no. 1, 2012, pp. 6–22.

[6] R. Barn, “Social media and protest - the Indian Spring?” Huffington Post, 2013, Retrieved: 10-07-2017. [Online]. Available: hhttp://www.huffingtonpost.co.uk/professor-ravinder-barn/india-social-media-and-protest_b_2430194.html

[7] S. Ahmed and K. Jaidka, “The common man: An examination of content creation and information dissemination on Twitter during the 2012 new delhi gang-rape protest,” in International Conference on Asian Digital Libraries. Springer, 2013, pp. 117–126.

[8] F.-M. Titzmann, “The voice of the youth, locating a new public sphere between street protest and digital discussion,” Studying Youth, Media and Gender in Post-Liberalisation India: Focus on and beyond the Delhi Gang Rape, vol. 6, 2014, p. 79.

[9] E. Morozov, “Iran: Downside to the "Twitter revolution",” Dissent, vol. 56, no. 4, 2009, pp. 10–14.

[10] A. M. Attia, N. Aziz, B. Friedman, and M. F. Elhusseiny, “Commentary: The impact of social networking tools on political change in Egypt’s "Revolution 2.0",” Electronic Commerce Research and Applications, vol. 10, no. 4, 2011, pp. 369–374.

[11] M. C. Stoddart and D. Tindall, “We’ve also become quite good friends: Environmentalists, social networks and social comparison in British Columbia, Canada,” Social Movement Studies, vol. 9, no. 3, 2010, pp. 253–271.

[12] E. Morozov, The net delusion: The dark side of Internet freedom. PublicAffairs, 2012.

[13] F. Rebillard and A. Touboul, “Promises unfulfilled? Journalism 2.0, user participation and editorial policy on newspaper websites,” Media, Culture & Society, vol. 32, no. 2, 2010, pp. 323–334.

[14] J. Go, “For a postcolonial sociology,” Theory and Society, vol. 42, no. 1, 2013, pp. 25–55.

[15] E. W. Said, Orientalism. Vintage, 1979.

[16] G. C. Spivak, Outside in the teaching machine. Routledge, 2012.

[17] M. G. Durham, “Scene of the crime: News discourse of rape in India and the geopolitics of sexual assault,” Feminist Media Studies, vol. 15, no. 2, 2015, pp. 175–191.

[18] P. Roychowdhury, “"The Delhi gang rape": The making of international causes,” Feminist studies, vol. 39, no. 1, 2013, pp. 282–292.

[19] P. Burnap, O. Rana, M. Williams, W. Housley, A. Edwards, J. Morgan, L. Sloan, and J. Conejero, “COSMOS: Towards an integrated and scalable service for analysing social media on demand,” International Journal of Parallel, Emergent and Distributed Systems, vol. 30, no. 2, 2015, pp. 80–100.

[20] N. Kourtellis, J. Finnis, P. Anderson, J. Blackburn, C. Borcea, and A. Iamnitchi, “Prometheus: User-controlled p2p social data management for socially-aware applications,” in Proceedings of the ACM/IFIP/USENIX 11th International Conference on Middleware. Springer-Verlag, 2010, pp. 212–231.

[21] N. Menon, “India’s Daughter and sexual violence: The issues at stake - Europe Solidaire Sans Frontieres,” 2015, Retrieved: 10-07-2017. [Online]. Available: http://www.europe-solidaire.org/spip.php?article34577

[22] K. Krishnan, “Rape, rape culture and the debate over India’s Daughter,” 2016, Retrieved: 10-07-2017. [Online]. Available: http://scroll.in/article/711369/rape-rape-culture-and-the-debate-over-indias-daughter

[23] V. Narayan, “Why the documentary India’s Daughter should be aired,” 2015, Retrieved: 10-07-2017. [Online]. Available: https://vinodnarayan.com/2015/03/04/should-the-documentary-indias-daughter-be-aired/

[24] S. Dev, “What does banning the documentary India’s Daughter accomplish exactly?” 2015, Retrieved: 10-07-2017. [Online]. Available: https://sonalidev.com/2015/03/04/what-does-banning-the-documentary-indias-daughter-accomplish-exactly/

[25] S. Krishnan, “India’s Daughter: When men come together, but women stand divided,” 2015, Retrieved: 10-07-2017. [Online]. Available: http://www.tarshi.net/inplainspeak/indias-daughter-when-men-come-together-but-women-stand-apart/

[26] I. Facts, “NLTK 3.2.4 documentation,” 2017, Retrieved: 10-07-2017. [Online]. Available: http://www.nltk.org

[27] A. Schulz, E. L. Menca, and B. Schmidt, “A rapid-prototyping framework for extracting small-scale incident-related information in microblogs: Application of multi-label classification on tweets,” Information Systems, vol. 57, 2016, pp. 88–110.

[28] C. J. Hutto and E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text,” in Eighth International AAAI Conference on Weblogs and Social Media, 2014.

[29] M. A. Smith, B. Shneiderman, N. Milic-Frayling, E. Mendes Rodrigues, V. Barash, C. Dunne, T. Capone, A. Perer, and E. Gleave, “Analyzing (social media) networks with NodeXL,” in Proceedings of the fourth international conference on Communities and technologies. ACM, 2009, pp. 255–264.

[30] M. Bastian, S. Heymann, M. Jacomy et al., “Gephi: an open source software for exploring and manipulating networks.” ICWSM, vol. 8, 2009, pp. 361–362.

[31] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” Journal of statistical mechanics: theory and experiment, vol. 2008, no. 10, 2008, p. P10008.

[32] E. A. Tiryakian, A. Gruzd, B. Wellman, and Y. Takhteyev, “Imagining Twitter as an imagined community,” American Behavioral Scientist, vol. 55, no. 10, 2011, pp. 1294–1318.

[33] P. Howlader and K. Sudeep, “Degree centrality, eigenvector centrality and the relation between them in Twitter,” in Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on. IEEE, 2016, pp. 678–682.

[34] W. H. Dutton, “The fifth estate emerging through the network of networks,” Prometheus, vol. 27, no. 1, 2009, pp. 1–15.

[35] D. Cohen, “BBC - letter from Danny Cohen to Sh. Rakesh Singh regarding concerns about Storyville - India’s Daughter - Media Centre,” 2015, Retrieved: 10-07-2017. [Online]. Available: http://www.bbc.co.uk/corporate2/mediacentre/statements/dc-letter-re-indias-daughter

[36] B. Mehra, “India’s Daughter: An indian daughter’s opinion,” 2014, Retrieved: 10-07-2017. [Online]. Available: http://swarajyamag.com/culture/indias-daughter-an-indian-daughters-opinion/

[37] A. Zubiaga, M. Liakata, R. Procter, G. W. S. Hoi, and P. Tolmie, “Analysing how people orient to and spread rumours in social media by looking at conversational threads,” PloS one, vol. 11, no. 3, 2016, p. e0150989.

[38] D. Murthy, “Digital ethnography an examination of the use of new technologies for social research,” Sociology, vol. 42, no. 5, 2008, pp. 837–855.

[39] L. Townsend and C. Wallace, “Social media research: A guide to ethics,” 2015, Retrieved: 10-07-2017. [Online]. Available: http://www.gla.ac.uk/media/media_487729_en.pdf

[40] J. Boorstin, “Facebook, Snapchat and Twitter played a bigger role than ever in the election,” 2017, Retrieved: 10-07-2017. [Online]. Available: http://www.cnbc.com/2016/11/08/facebook-snapchat-and-twitter-played-a-bigger-role-than-ever-in-the-election.html

# Modeling User Behavior in Social Media with Complex Agents

Jan Ole Berndt, Stephanie C. Rodermund, Fabian Lorig, Ingo J. Timm

TriLabS @ CIRT, Business Informatics 1
Trier University
54296 Trier, Germany
Email: [berndt,s4strode,lorigf,itimm]@uni-trier.de

***Abstract*—Social media like Facebook, Twitter, or Google+ have become predominant means of communication. However, their distributed structure and dynamic interaction processes make it difficult to analyze and understand that communication. Thus, we propose agent-based modeling and simulation of user behavior for analyzing communication dynamics in social media. We develop an agent decision-making method that models motivations of media users and their impact on behavior by means of social actor types. Moreover, we apply this model to Twitter communication accompanying a German television program. Our evaluation shows that different actor constellations within a population of agents drastically impact the dynamics of this communication.**

***Keywords–Agent-Based Modeling; Social Actor Types; Social Media Analysis; User Behavior; Social Simulation.***

## I. Introduction

One of the most noticeable advances of this century is the omnipresence of information and communications technology. The establishment of computer systems in our daily life and the connection of private households to the Internet has initiated and still promotes the *digital revolution* [1]. In particular, *social media* like Facebook, Twitter, or Google+ have become predominant means of communication for both private and professional users. They are widely used for various purposes, ranging from casual smalltalk to commercial marketing campaigns and the shaping of political opinion [2] [3].

Understanding these communication processes is important in both commerce and politics to derive communication strategies for social media. For example, marketing campaigns can reach a vast target audience through viral communication dynamics [4]. However, if the same dynamics distributes negative opinions, emerging mass criticism can endanger a company's commercial success. Therefore, it is crucial to anticipate likely reactions of social media users to such campaigns to avoid unintended effects or to develop appropriate counter strategies to those effects [5].

Nonetheless, the inherent distribution of social media and the dynamics of user interactions therein make it difficult to analyze and understand that kind of communication. Thus, manual analysis has been complemented with computational linguistics, data mining, and simulation methods [6]. These methods help recognize conversation topics, discern user communities, and model information diffusion in social networks.

Especially agent-based social simulations [7] are a promising technique for understanding complex dynamics of interrelated communication activities. They model behaviors of humans by means of artificial agents in order to explore the effects of different social actor constellations and various situations in an experimental environment. For instance, viral dynamics of mass phenomena in social media like the *harlem shake* [8] can be reproduced by using artificial agents for representing media users [9]. Each agent can react to other agents' communication activities in a simulated media environment. This interaction leads to complex dynamics. Exploring various user constellations and agent decisions in a controlled experiment helps understand these dynamics in real world social media.

However, agent-based simulation for social media analysis requires a model of user motivations and resulting behaviors to yield realistic results. Agents must be complex enough to explain *why* particular communication processes emerge and which effects potential reactions to them will provoke. Thus, in this paper, we develop an agent-based model of user behavior for analyzing communication dynamics in social media. This is a first step toward a simulation-based decision-support method for developing and testing social media communication strategies as proposed by Berndt et al. [5].

The paper is structured as follows. Section II provides an overview of the foundations of social media analysis, social actor theory, and agent-based modeling as a technique for dynamic analysis. Subsequently, Section III describes our concept of complex agents for modeling user behavior. This concept covers individual social actors, their respective decision-making, as well as populations of media users. Section IV applies that concept to communication processes on Twitter which accompany a German television program. In Section V, we evaluate our model by simulating user behaviors in that scenario. Finally, Section VI concludes on our findings and gives an outlook on possible future work.

## II. Foundations

To analyze, model, and simulate user behavior in social media, it is necessary to understand communication processes within those media. These processes depend on the underlying platforms that structure possible communication, the observable communicative activities, as well as the social actors performing these activities. Thus, the following sections discuss approaches and theories for analyzing and modeling these aspects. In addition, we give an overview of the state of research in agent-based modeling of human behavior to provide a foundation for our approach to user behavior analysis.

### A. Social Media

Social media structure communication processes by providing options to their users to connect with each other. In terms of graph theory, such a structure can be described by a set of users (nodes) and relationships between the users (edges)

 ISBN: 978-1-61208-578-4 

[10]. Graphs can be unidirectional, defining the direction of the relationship, or bidirectional, connecting two nodes without providing information regarding that direction.

For instance, the online social network Twitter can be modeled as a directed graph. In contrast to most other platforms, which consist of bidirectional relationships between users, a distinction between *followers* and *followees* is made on Twitter. That is, a user actively and voluntarily decides which other users to *follow* for receiving their status updates. Following another Twitter participant makes the following user become a *follower*. However, a *followee*, i.e., the user being followed, does not need to follow his or her followers.

When analyzing the structure of social media, a typical task is to identify and assess the importance of the most influential users by means of centrality measures [11]. The *degree* of centrality corresponds to the total number of edges a node has. Hence, it is a measure of a node's interconnectedness in a graph. Nodes having a high *degree* (compared to other nodes) act as hubs for information diffusion within a social network.

By contrast, a graph's *density* denotes the interconnectedness of an entire network. It is used for comparing different network structures and their impact on information propagation. The *density* is defined by the ratio of the number of existing edges and the maximum number of edges in case every pair of nodes would be connected by an edge (complete graph).

### B. Communication

Human communication can be considered as a sequence of actions by individuals, where the behavior of a sender influences the behavior of a receiver [12]. The sender uses a set of characters to encode a message, which is transmitted using an information medium. The receiver uses an own set of characters to decode and interpret the message and returns a feedback using the same mechanism [13]. The formulation and transmission of messages by the sender as well as the corresponding reaction by the receiver form the communicative activities available to users of social media.

However, the shift of communication into technical media is accompanied by a loss of information. The transmission of messages is ensured, yet, the receiver does not know whether a message was interpreted correctly. On Twitter, communication results can only be returned by replying to a Tweet using another Tweet. Consequently, conversations are formed as sequences of messages which refer to or forward previous ones [14]. To that end, Twitter provides mechanisms for replying to other tweets and for addressing a tweet to a certain person. Using the @-symbol followed by the name of a user or by putting the prefix *"RT"* (retweet) at the beginning of a tweet, the identification of dialogs or conversations is supported.

In addition to the structuring of dialogs, Twitter users can use another operator for classifying the content of a message. The content provides information about the intention as well as the context of communication. On Twitter, the #-symbol (hashtag) is used for categorizing messages and for marking keywords. This simplifies filtering Tweets according to certain topics, which makes this kind of communication easily accessible to media studies and communication research. In fact, Twitter has been widely used for conducting studies of certain subjects or events, e.g., spread of news and criticism [15] [6], the activity of diseases [16], or political communication [3].

### C. Social Actors

Communication is inherently social. In fact, sociality can be considered to consist entirely of communication [17]. Social systems emerge from interconnected communicative activities being selected by social actors. Those actors are influenced by an observed social situation. They decide about their reactions to that situation. This results in observable behaviors that lead to a new situation in effect (Figure 1). For example, a user can observe an ongoing conversation about a specific topic (1). She may decide to utter a controversial opinion about that topic (2). Her utterance becomes observable to other users in the form of her respective Tweet (3). This changes the conversation and provokes further reactions. Thus, the conversation on the macro-social level (4) both influences individual behaviors and emerges from them on the micro-social level.

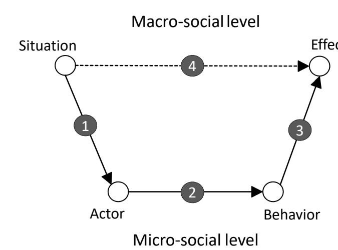


Figure 1. Emergence of macro-social effects from micro-social behavior [18].

There are several analyses of user behavior in social media available. For instance, activity frequencies on Twitter (i.e., Tweets, Responses, Retweets) have been related to user attributes and traits such as gender, age, region and political opinion [19]. While such an analysis reveals *how* social media users interact with each other, it cannot explain *why* they do it. To answer that question, other studies cover motivations for communication. These motivations can be categorized into groups like *smalltalk*, *entertainment*, or *information and news sharing* [20]. Additionally, they can be derived from psychological personality traits [9] [21]. Such approaches provide valuable insights into the decision-making process of social actors in diverse situations ranging from casual comments on a television series [22] to crisis communication [23].

In addition to social media specific and psychologically founded motivational categories, there are also theories of actor behavior in sociology. Sociologists distinguish between four basic social actor types which differ in their behavior [24]. Firstly, the *homo economicus* is a rational decision-maker who strives to maximize her personal utility. Such an actor attempts to reach personal goals as efficiently as possible. Secondly, the *homo sociologicus* obeys social norms and obligations. This actor type tries to conform with expectations to avoid negative sanctions. Thirdly, the *emotional man* is driven by uncontrollable emotions such as love, anger, respect, or disgust. This leads to affective behavior in response to, e.g., unfulfilled expectations [25]. Finally, the *identity keeper* has the goal to establish and maintain a desired social role. Such an actor seeks social acknowledgment by provoking positive reactions toward stereotypical behaviors. In the remainder of this paper, we will show how these actor types can be applied to agent-based modeling of user behavior in social media.

 ISBN: 978-1-61208-578-4 

*D. Related Work: Agent-Based Modeling of Human Behavior*

As discussed in the preceding section, communication processes in social media emerge from individual activities of the participating users. For investigating emergent phenomena, agent-based modeling has been established as a standard means. Artificial agents are capable of decision-making, communication, and goal-directed behavior [26]. By modeling real world actors as software agents, individual behavior and anticipation of behavior on the micro level can be simulated resulting in emergent effects on a macro level [27] [28]. In terms of social sciences, using such actor models for simulation studies is referred to as agent-based social simulation [7].

The majority of agent-based models in social media analysis is concerned with *information propagation*. They aim at identifying a group of users which can propagate information, i.e., a message, to as many users as possible [29]. The users are frequently modeled as agents being connected by neighborhood relations in cellular automata [30] or general network graphs [31]. These agents often have particular behavioral rules that fire if a certain activation threshold is reached. Such a threshold denotes the required strength of influence (e.g., a number of received messages) on an agent until it becomes active itself. This method is particularly relevant for planning advertising strategies since viral marketing campaigns make use of information propagation effects [32] [4].

While threshold models are usually investigated by means of simulation studies, there are also *analytical approaches* to agent-based modeling of opinion formation. These focus on the interactions among agents which lead to the diffusion and adoption of opinions in a process of compromising [33]. They model these interactions by means of thermodynamics [34] or the kinetic theory of gases [35]. These methods describe the emergence of macro-social phenomena from micro-social interactions using differential equations. This allows for analyzing the resulting opinion dynamics mathematically.

However, there is a discrepancy between these threshold and analytical models on the one hand, and the mentioned sociological perspectives on decision-making on the other. While these methods describe *how* opinion and communication dynamics occur in agent-based social simulations, they lack the descriptive power to analyze *why* this happens. That is, they focus on the dynamics between interacting agents and treat the agent population as a homogeneous mass. For instance, in kinetic theory, gas molecules behave solely according to their current states and their mutual influences without having individual habits. The same holds for cellular automata in which all cells, i.e., agents, are usually homogeneous and strongly restricted in their neighborhood relations. As a result, the discussed approaches largely disregard modeling individual motivations for decision-making such as described by social actor types.

For utilizing agent-based social simulation to understand human behavior and to develop communication strategies, it is necessary to apply more elaborate agent decision approaches. Agents must have individual motivations to allow for analyzing who participates in communication processes for which reason [5]. Since, in social media, different users react differently to the same message, this should also be the case for artificial agents in a simulation model. In fact, a wide range of agent decision-making architectures based on philosophy, psychology and cognitive science is readily available [36]. In addition, sociological theory and agent-based modeling have been combined in the interdisciplinary field of *socionics* [37]. In that context, the described social actor types can be utilized to explain social behavior in an agent-based simulation.

Dittrich and Kron model social characters by means of actor types and combinations between these types [24]. They simulate the so-called "bystander dilemma" in which persons must decide whether or not to help a victim of physical violence. In their model, agents implementing the *homo sociologicus* and *identity keeper* roles feel obliged to help while *homo economicus* and *emotional man* flee the situation. Combining these dispositions on both an individual and on a population level leads to complex macro-social behaviors. This makes that approach a promising candidate for a transfer to modeling user behavior in social media as described in the following section.

## III. CONCEPT: MODELING USER BEHAVIOR

In this section, we adapt the agent-based decision-making approach by Dittrich and Kron [24] to modeling communicative user behavior in social media. That is, we model the selection of messages about a specific topic to be published on a social media platform within a limited time frame [38].

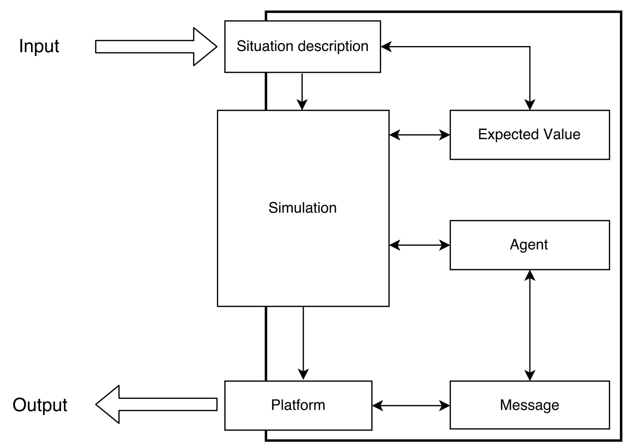


Figure 2. Structure of the user modeling and simulation concept.

Our modeling and simulation concept is structured as depicted in Figure 2. Each decision-making situation receives an input of one or more keywords to describe that situation (e.g., a list of hashtags or abstract topic description). The respective output consists of messages being published at the social network platform by the population of agents. In order to produce that output, each agent observes the situation and calculates expected values for its potential reactions according to its respective social actor type and depending on the activities of other agents. It then selects its next message (or chooses not to publish any message) with respect to these expected values. The following sections describe the actor types, their combinations, and the resulting agent populations.

### A. Social Actor Types and Decision-Making

Besides the current situation, its social actor type determines an agent's decision-making. To that end, we model each type by means of a function $EV$ that returns an expected value for each available activity option. For a *homo economicus*, this amounts to a standard utility function. Contrastingly, a *homo* *sociologicus* prefers socially adequate behaviors over controversial actions. Such an agent makes its behavior dependent on contributions to a conversation by other agents. In addition, while the *identity keeper* has a genuine desire to further any kind of discussion, the *emotional man* only becomes active when being emotionally affected by the situation.

All of the expected value functions should cover the same range of values to make them comparable with each other. That range depends on the number of available activity options and their effects in a particular application scenario. Each option can either have a positive, neutral, or negative effect on an agent's goals. For instance, a scenario with five possible messages can be encoded through the following set of values: $\{-1, 0, 1, 2, 3\}$. In this case, a message is either detrimental to an agent's goals (-1), it can be neutral towards them (0), or it furthers its motivations to different extents (1–3). Then, the agent can select its actions as follows.

$$\arg\max_a EV_i(s, a) \tag{1}$$

Each actor type $i$ ($i \in \{$*homo economicus*, *homo sociologicus*, *emotional man*, *identity keeper*$\}$) maximizes its expected value for all available actions $a$ in the current situation $s$. If there are several options with the same value, an agent decides randomly among them. This results in a specific message (i.e., Tweet) being selected and published at the simulated social network platform for all other agents to observe.

Using the described value maximization approach to select a message to be communicated leads to a restriction in the amount of behavioral randomness. This is especially useful for evaluating the sensitivity of the resulting emergent effects on the population level to the agent population. Different compositions of agents within a population will lead to different interactions with low variance.

In order to increase the variance of agent behaviors, fluctuating populations can be introduced. Alternatively, a random selection of messages, weighted by their respective expected values, can be introduced. This will then increase the randomness on an individual instead of the population level. However, adding this stochasticity decreases the explanatory impact of modeling social actor types because their respective motivations become less pronounced in the selected communication activities.

### B. Actor Type Combinations and Populations

According to the preceding decision-making model, each agent can implement one of the four available actor types. However, these are only prototypical examples for categorizing motivations. In fact, an actor's social disposition will often be more adequately described by a mixture of several basic motivations [24]. Consequently, we allow for combinations of actor types within individual agents to represent that phenomenon.

For mixing several actor types, each agent is defined by four weights $w_i$, one for each actor type $i$, with $\sum_i w_i = 1$. The weights denote the ratio with which those types contribute to its decision-making. Then, an agent with mixed types selects its activities by maximizing the weighted sum of the respective expected values (with a randomized selection in case of several maxima).

$$\arg\max_a \sum_i EV_i(s, a)\ w_i \tag{2}$$

In addition to combining actor types within an individual agent, it is also possible to mix different agents within the overall agent population. That is, a population can either consist of homogeneous agents that all implement the same actor type combination, or it can comprise different agents. Homogeneous populations are particularly useful for model validation and calibration. They make the effects of different value functions easily observable and adjustable. Contrastingly, heterogeneous populations are more realistic. They lead to complex interaction dynamics which are necessary for replicating and explaining user behaviors in social media as described in the following sections.

## IV. Application: Agent-Based Analysis of Social Media Communication

In this section, we apply our agent-based modeling concept to an analysis of user behavior in communication processes on Twitter. In particular, we model live-tweeting behavior during an episode of the German television series "Tatort" (meaning *crime scene*). Running since 1970, "Tatort" is the most popular German TV series, which attracts a broad audience across all social groups, genders, and ages. We use a dataset of Tweets about the episode "Alle meine Jungs" (*all my boys*), of 18 May 2014. The dataset contains eight distinct phases of very high or very low Twitter activity which correspond to specific scenes of the episode. These scenes provide the situation for the agents in our model to react to. Each of them is described by one or more out of five attributes as shown in Table I.

TABLE I. SITUATION DESCRIPTIONS.

| Scene | Description | Scene | Description |
|---|---|---|---|
| 0 | thrilling | 4 | funny |
| 1 | funny, music-related | 5 | thrilling, emotional |
| 2 | funny, music-related | 6 | thrilling |
| 3 | funny, music-related | 7 | judgmental |

In our model, the agents can act repeatedly during each scene. At the beginning of a scene, they base their actions only on the respective description; subsequently, they can also react to other agents' Tweets. Thus, a dynamic communication system emerges from these interrelated activities. In the following, we describe the available actions and the decision-making of the four actor types in these situations.

### A. Agent Activity Options

The Tweets in our dataset can be classified by their sentiment and tonality along two different dimensions. They are either positive or negative and they are either joking or not joking (i.e., serious). The possible combinations of these categories result in four different message types available to the agents. However, since not all users reply to every message, an agent also has the option not to tweet. Nevertheless, it can still decide to participate in the conversation about the current scene at a later time after observing Tweets by other agents. This results in the following five activity options for the agents.

1) No Tweet
2) Tweet – positive – joking
3) Tweet – positive – not joking
4) Tweet – negative – joking
5) Tweet – negative – not joking

 ISBN: 978-1-61208-578-4 

Which of these options an agent selects at which time depends on its underlying combination of actor types, as well as on the activities of other agents as described in the following.

### B. Agent Decision-Making

In our application example, the actor types defining the agents' decision-making represent typical behavioral roles and motivations in social media communication. These include the maximization of publicity, a desire for serious discussion, the expression of anger, as well as genuine content production. These motivations are represented by the *homo economicus*, *homo sociologicus*, *emotional man*, and *identity keeper*, respectively. For all actor types, we evaluate the available activity options with respect to those motivations in each situation in order to identify expected values for the agents' decisions [39]. Table II summarizes the criteria and values for that evaluation.

TABLE II. DECISION-MAKING BY SOCIAL ACTOR TYPES.

| Homo Economicus | Homo Sociologicus | Emotional Man | Identity Keeper |
|---|---|---|---|
| No Tweet (0) | Must (3) | Unchanged (0) | Strengthened (3) |
| Utility function (0 to 3) | Should (2) | Increased (-1) | Weakened (-1) |
| Conversation size threshold (-1) | Can (1) | Decreased (2) | |
| | Should not (-1) | Strongly decreased (3) | |

In social media communication, a *homo economicus* agent attempts to maximize the impact of its contributions on the conversation. Such an agent gains the highest utility by provoking agreement with as many other agents as possible. Thus, its underlying utility function anticipates probable majority opinions. Actions supporting these are rated higher than less popular or even controversial contributions according to the distribution of actions in the original dataset. This agent type will maintain its ratings during an actual conversation regardless of other agents' behaviors.

In addition, we use a threshold of a minimal number of Tweets by other agents for this type of agent to become active itself. This threshold equals to the mean number of Tweets across all scenes (24 in the dataset). Until the threshold is reached, an agent will not participate in the conversation, leaving its utility unchanged. Thus, the *homo economicus* represents a casual media user who only joins ongoing conversations to represent common sense opinions shared by the expected majority of recipients.

Contrastingly, a *homo sociologicus* agent rates the available actions according to general social norms as well as other agents' behaviors. With respect to the scene description, its expected value function evaluates these options by their perceived strength of obligation. For instance, an agent *should not* joke about an emotional scene. However, if the majority of other agents has deviated from such norms before, the *homo sociologicus* will mimic these previously observed activities in order to gain acceptance by other agents. Hence, that type of agent represents a both morally concerned and opportunistic user who joins the dominant group as soon as one emerges. This behavior is typical, e.g., in massive online protests [6].

The *emotional man*, on the other hand, represents an outright dissatisfied and angry user. Such an agent strives to express that anger. This leads to predominantly negative and sometimes sarcastic (i.e., joking) contributions. By publishing negative Tweets, the agent decreases its anger until it no longer feels the need to communicate. Consequently, that behavior produces isolated criticism without any intention of engaging in an actual discussion.

Finally, the *identity keeper* is a genuine content producer. This type of agent has the goal of bringing forward any kind of discussion in order to maintain its participation in it. That is, the agent can strengthen its identity by providing arguments for other agents to react to. For that purpose, any kind of Tweet can be appropriate, especially controversial ones if they provoke reactions. Only remaining inactive weakens that identity. As a result, the *identity keeper* represents a user who enjoys a conversation for the sake of the conversation and who ensures a certain diversity of perspectives on the discussed topic.

By combining the described actor type models within individual agents, it is possible to represent mixed motivations and to implement a wide variety of decision behaviors. Moreover, heterogeneous populations of different agent types will lead to complex interactions of these behaviors. The following section evaluates these effects.

## V. EVALUATION: SIMULATION OF USER BEHAVIOR

As a proof of concept for our agent-based modeling approach, we have implemented the aforementioned agent types and decision-making algorithms in a *JAVA* program. In the following, we use that program to simulate user behaviors emerging from different populations of various agents. Such a simulation gives a first impression of the range of effects that the model can (re)-produce. In particular, it allows for analyzing the interplay between several actor types on both the individual and the population level.

In our simulation, we compare two different settings. The first one consists of a homogeneous agent population with mixed actor types. That is, each agent combines all four types with equal weights. By contrast, the second setting comprises a heterogeneous agent population in which every agent implements one of the four basic actor types. Throughout the population, these agents are uniformly distributed. They communicate about all eight scenes. Their respective activity choices depend on the situation description for those scenes as well as on the previous actions of other agents.

For both settings, the population size is set to 164 agents that can join the conversation in each scene (as in the real world dataset). This number is relevant as long as the *homo economicus* uses a fixed conversation size threshold. The more agents there are, the sooner will a *homo economicus* impact the communication dynamics. While the threshold can be scaled up or down according to the population size, we use the realistic one to enable comparisons of our simulation results with that dataset in future studies.

Figure 3 shows the arithmetic mean of our evaluation results together with the respective standard deviations out of 100 simulation repetitions (except for the "No Tweet" option). For the homogeneous population, the results show a majority of negative not joking Tweets. This is due to the fact that both *identity keeper* and *homo economicus* consider this activity as adequate. Moreover, the *emotional man* favors it over all others. Combining these within the agents leads to the observed uniformity which even becomes amplified as the *homo sociologicus* imitates dominant behaviors. Only scene 4 leads to negative as well as positive Tweets. That scene is described as being funny. Hence, the positive actions favored

 ISBN: 978-1-61208-578-4 

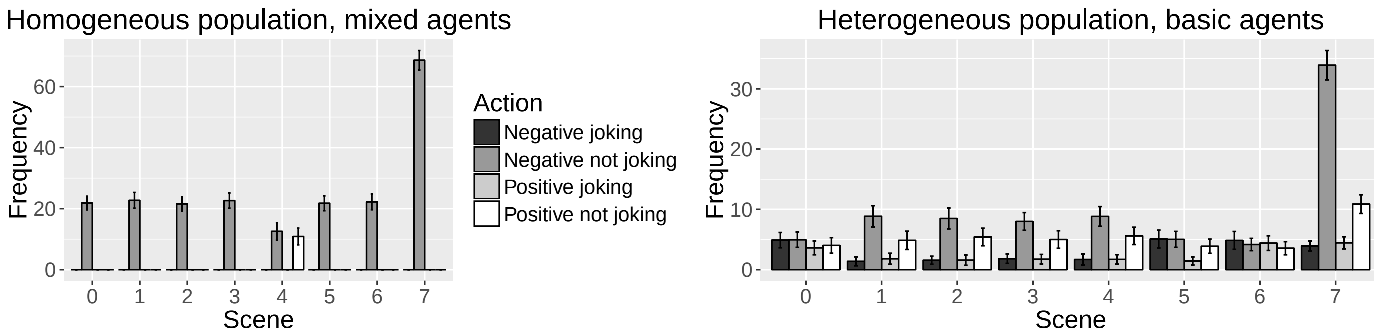


Figure 3. Activity frequencies of a homogeneous population of mixed actors (left) and a heterogeneous population of basic actor types (right).

by all other actor types override the negative option selected by the *emotional man*.

Contrastingly, the heterogeneous population leads to more diverse behavior. In that case, negative Tweets are still prevalent for most scenes. This is caused by the same effects as described: The *homo sociologicus* amplifies the behavior being initially driven by the other actor types, particularly the *emotional man*. However, since these agent types act simultaneously in a mixed population, all other actions are also observable. This leads to realistic effects, such as decisions not to tweet at all in scenes being described as thrilling.

Overall, these results show that the combination of actor types both within individual agents and their mixing in heterogeneous populations drastically impacts the emergent dynamics of simulated social media communication. This demonstrates that modeling motivations of individual agents can produce behavioral heterogeneity, which other models have to introduce artificially, e.g., by means of random noise [35]. In contrast to those approaches, we can directly control which type of agent reacts to which particular communicative situation in what manner. The composition of an agent population then models the affinity or aversion of a user group in social media to certain topics, opinions, and communication styles. Hence, we conclude that our model adds this composition of populations as an important variable to existing methods for studying information diffusion in social simulations.

However, it is important to select and calibrate the agent types carefully for such a simulation to yield meaningful results for understanding user behavior. To that end, it is necessary to analyze available real world data and identify typical activity patterns [6]. Then, potential underlying motivations can be derived from those observations in order to define the required actor types and their combinations [9] [21]. With this work, we have shown how such actor types can be modeled for exploring user behavior in agent-based social simulations.

## VI. Conclusions and Future Work

In this paper, we have developed an agent-based model of user behavior in social media. This model facilitates dynamic analyses of complex communication processes which are difficult to assess by means of conventional approaches. In such a context, agent-based social simulations allow for experimentally exploring emergent behaviors [6].

Our model focuses primarily on the decision-making of social actors communicating about a specific topic. This is in contrast to existing work on information diffusion, which analyzes the impacts of social network structures on the spread of messages. Instead, we have modeled motivational causes for user behaviors by utilizing complex agents based on sociological theory. To that end, we have presented a general concept for representing and combining four different actor types in agent-based social simulations. In addition, we have applied this concept to model and analyze Twitter communication about a German television program. Our evaluation shows that particular combinations of different motivations either within individual agents or across an entire population drastically impact communication dynamics. Therefore, we conclude that it is crucial to consider these motivations carefully in order to realistically model and explain user behavior in social media.

While our model provides a promising first step to agent-based simulations of social media usage, there are several extensions we consider for future work. Firstly, we are working on calibrating the model to accurately imitate the user interactions observed in our real world example. This will provide insight into the achievable realism when combining the four basic actor types into complex agents and populations. As a first result, we have already demonstrated that our model is indeed capable of reproducing the communication activities of real world users [39].

Secondly, it would be interesting to integrate the agent decision method with existing information diffusion approaches [29]. This will complement those methods with motivational aspects of *why* information is spread within a social network. In that context, the population composition will provide an additional variable which impacts communication dynamics. The various actor types can then produce behavioral heterogeneity on a more detailed and explanatory deeper level than the addition of abstract random noise to an equational modeling approach [35].

Finally, it will also be necessary to model the activity options for the agents in more detail. This covers particularly the message contents. In order to simulate, e.g., the shaping of opinions in political discourses, a classification of communication contents and their impact on the interaction is required. To achieve this, we plan to utilize content modeling and annotation techniques from media and communication studies [40] for encoding discourses in agent-based social simulations.

## Acknowledgments

We would like to thank Carla Schmidt, Christof Barth, and Hans-Jürgen Bucher for providing us with real world data and a media studies perspective on our application example.

 ISBN: 978-1-61208-578-4 

## References


[1] W. R. Neuman, Y. Park, and E. Panek, "Tracking the Flow of Information into the Home: An Empirical Assessment of the Digital Revolution in the United States, 1960-2005," International Journal of Communication, vol. 6, 2012, pp. 1022–1041.

[2] J. Kirby, Connected marketing: the viral, buzz and word of mouth revolution. Amsterdam: Butterworth-Heinemann, 2010.

[3] A. Maireder and S. Schlögl, "24 hours of an# outcry: The networked publics of a socio-political debate," European Journal of Communication, 2014, pp. 1–16.

[4] D. Kempe, J. M. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network." Theory of Computing, vol. 11, no. 4, 2015, pp. 105–147.

[5] J. O. Berndt, F. Lorig, I. J. Timm, C. Barth, and H.-J. Bucher, "A systematic approach to agent-based dynamic analysis of social media communication," International Journal On Advances in Internet Technology, vol. 10, no. 1&2, to appear 2017.

[6] I. J. Timm, J. O. Berndt, F. Lorig, C. Barth, and H.-J. Bucher, "Dynamic analysis of communication processes using twitter data," in 2nd International Conference on Human and Social Analytics (HUSO 2016). IARIA, 2016, pp. 14–22.

[7] P. Davidsson, "Agent based social simulation: A computer science view," Journal of Artificial Societies and Social Simulation, vol. 5, no. 1, 2002.

[8] J. Bollen, H. Mao, and A. Pepe, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena." in 5th International AAAI Conference on Weblogs and Social Media, 2011, pp. 450–453.

[9] F. Lorig and I. J. Timm, "How to model the human factor for agent-based simulation in social media analysis?" in 2014 ADS Symposium (part of SpringSim multiconference). SCS, 2014, p. 12.

[10] F. Vega-Redondo, Complex Social Networks. Cambridge University Press Cambridge, MA, 2007.

[11] L. C. Freeman, "Centrality in social networks conceptual clarification," Social networks, vol. 1, no. 3, 1978, pp. 215–239.

[12] C. R. Berger, "Interpersonal communication," The International Encyclopedia of Communication, 2008.

[13] C. E. Shannon, "A mathematical theory of communication," ACM SIGMOBILE Mobile Computing and Communications Review, vol. 5, no. 1, 2001, pp. 3–55.

[14] D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," in 43rd Hawaii International Conference on System Sciences (HICSS). IEEE, 2010, pp. 1–10.

[15] K. Lerman and R. Ghosh, "Information contagion: An empirical study of the spread of news on digg and twitter social networks." ICWSM, vol. 10, 2010, pp. 90–97.

[16] A. Signorini, A. M. Segre, and P. M. Polgreen, "The use of twitter to track levels of disease activity and public concern in the us during the influenza a h1n1 pandemic," PloS one, vol. 6, no. 5, 2011, p. e19467.

[17] N. Luhmann, Social Systems. Stanford, USA: Stanford University Press, 1995.

[18] P. Hedström and P. Ylikoski, "Causal mechanisms in the social sciences," Annual Review of Sociology, vol. 36, 2010, pp. 49–67.

[19] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta, "Classifying latent user attributes in twitter," in 2nd international workshop on Search and mining user-generated contents. ACM, 2010, pp. 37–44.

[20] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: understanding microblogging usage and communities," in 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis. ACM, 2007, pp. 56–65.

[21] S. Vandenhoven and O. De Clercq, "What does the bird say? exploring the link between personality and language use in dutch tweets," in 2nd International Conference on Human and Social Analytics (HUSO 2016). IARIA, 2016, pp. 38–42.

[22] S. Schirra, H. Sun, and F. Bentley, "Together alone: motivations for live-tweeting a television series," in 32nd annual ACM conference on Human factors in computing systems. ACM, 2014, pp. 2441–2450.

[23] V. Hoste, C. Van Hee, and K. Poels, "Towards a framework for the automatic detection of crisis emotions on social media : a corpus analysis of the tweets posted after the crash of germanwings flight 9525," in 2nd International Conference on Human and Social Analytics (HUSO 2016), 2016, pp. 29–32.

[24] P. Dittrich and T. Kron, "Complex reflexive agents as models of social actors," in Proceedings of the SICE Workshop on Artificial Society/Organization/Economy, ser. Meeting of Systems Engineering, vol. 25, Tokyo, 2002, pp. 79–88.

[25] H. Flam, "Emotionalman': I. the emotionalman'and the problem of collective action," International Sociology, vol. 5, no. 1, 1990, pp. 39–56.

[26] M. Wooldridge and N. R. Jennings, "Agent theories, architectures, and languages: a survey," in International Workshop on Agent Theories, Architectures, and Languages. Springer, 1994, pp. 1–39.

[27] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," Proceedings of the National Academy of Sciences, vol. 99, no. suppl 3, 2002, pp. 7280–7287.

[28] J. O. Berndt and O. Herzog, "Anticipatory behavior of software agents in self-organizing negotiations," in Anticipation Across Disciplines. Springer, 2016, pp. 231–253.

[29] C. Zhang, J. Sun, and K. Wang, "Information propagation in microblog networks," in 2013 IEEE/ACM international conference on advances in social networks analysis and mining. ACM, 2013, pp. 190–196.

[30] M. S. and B. F., "A stochastic model of self-stabilizing cellular automata for consensus formation," in 15th Workshop Dagli Oggetti agli Agenti (WOA 2014), Catania, 2014.

[31] A. Tsang and K. Larson, "Opinion dynamics of skeptical agents," in International conference on Autonomous agents and multi-agent systems (AAMAS 2014), 2014, pp. 277–284.

[32] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009, pp. 199–208.

[33] S. Monica and F. Bergenti, "Opinion dynamics in multi-agent systems: selected analytic models and verifying simulations," Computational and Mathematical Organization Theory, 2016, pp. 1–28.

[34] F. Schweitzer and J. A. Hołyst, "Modelling collective opinion formation by means of active brownian particles," The European Physical Journal B-Condensed Matter and Complex Systems, vol. 15, no. 4, 2000, pp. 723–732.

[35] S. Monica and F. Bergenti, "An analytic study of opinion dynamics in multi-agent systems with additive random noise," in AI* IA 2016 Advances in Artificial Intelligence. Springer, 2016, pp. 105–117.

[36] T. Balke and N. Gilbert, "How do agents make decisions? a survey," Journal of Artificial Societies and Social Simulation, vol. 17, no. 4, 2014.

[37] K. Fischer, M. Florian, and T. Malsch, Socionics: scalability of complex social systems. Berlin: Springer, 2005.

[38] R. Belkaroui, R. Faiz, and A. Elkhlifi, "Conversation analysis on social networking sites," in 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS). IEEE, 2014, pp. 172–178.

[39] S. C. Rodermund, F. Lorig, J. O. Berndt, and I. J. Timm, "An agent architecture for simulating communication dynamics in social media," in 15th German Conference on Multiagent System Technologies (MATES 2017). Berlin: Springer, to apper 2017.

[40] K. A. Neuendorf, The content analysis guidebook. Sage, 2016.

# In the Eye of the Beer-Holder.

# Lexical Descriptors of Aroma and Taste Sensations in Beer Reviews

Els Lefever, Liesbeth Allein and Gilles Jacobs
LT$^3$, Language and Translation Technology Team
Email: els.lefever@ugent.be, liesbeth.allein@ugent,be, gillesm.jacobs@ugent.be

***Abstract*—Western languages do not dispose of a well elaborated vocabulary for describing smell and flavour sensations. We investigate whether beer experts share a common vocabulary to describe beer properties. We collected an English text corpus of beer reviews and analyzed the lexical descriptors used by beer-tasting experts to describe aromas and flavours. The informativeness of beer reviews was investigated by running a machine learning experiment for predicting the colour of a beer based on the review text. This preliminary experiment shows promising results, with average accuracy figures of about 60% for automatic beer colour prediction. Our experimental results show that beer experts share a common vocabulary to describe beer characteristics in a consistent way, allowing to automatically predict beer properties based on the review text.**

***Keywords–Corpus linguistics; Natural language processing; classification; beer reviews***

## I. Introduction

It appears that people in the western world are not very good at describing smells and flavours. Research has shown that western languages have few words to describe smells and flavours, in contrast to visual phenomena, for which we dispose of a well elaborated vocabulary [1]. The description of taste, smell, and sight remains evaluative and non-specific for non-experts as [2] has shown in studies conducted with wine experts, coffee experts, and novices. Both coffee and wine experts tend to use more specific source-based terms (metaphors or *it smells like* + source), while novices use more evaluative terms (e.g., *nice, bad, good*). Consequently, we can assume that reviews written by beer experts should contain more source-based descriptions and fewer evaluative terms in order to describe smell, taste, and sight.

Previous research has shown that the perception of foods and drinks depend on both the visual and orthonasal sensory inputs, especially before the tasting [3]. The colour and look of a drink influence both the perception of smell and the way the taster describes his/her perceptions [4]. Furthermore, people link certain smells with certain colours [5]. It can therefore be concluded that perceptions of sight are closely related to perceptions of smell. Consequently, certain sight descriptions in the reviews will often be accompanied by the same smell descriptions and certain smell descriptions will be accompanied by the same sight descriptions. For this reason, an automatic prediction system could predict missing sight properties based on the smell descriptions in the reviews it is often accompanied by, and vice versa. For example, "gold" could refer to the colour of a beer, and in the corpus, reviews about light-coloured beer often contain the words *herbs*, *spicy*, and *butter* to describe the aroma of that beer. Therefore, an automatic prediction system could predict the colour of the beer in a review, which lacks sight or colour descriptions, based solely on the aroma description. Such a system could tell if a beer is gold even though the reviewer does not mention any colour or sight property.

Not only descriptions of smell and sight, but also descriptions of smell and taste/flavour are claimed to be closely connected. According to [6], the flavour of food is described by both gustatory and olfactory stimuli. There are two olfactory stimuli: orthonasal smell and retronasal smell, which are respectively the smell we sniff before the food or drink is tasted and the smell that is pulsed out after the food or drink had been swallowed. It is the combination of retronasal aromas and gustatory cues that defines a flavour and leads to descriptions such as *fruity* and *malty* [6]. In [7], the authors consider odour-taste synaesthesia (smelling tastes/tasting smells) a factor of the link between smell and taste. When people smell an odour, they recognize the smell as a taste and describe it as such, e.g., something *smells sweet*. This is also due to the co-occurrence of retronasal odour simulation and oral stimulation and the result of a unitary perception [7].

For this research, we have compiled a corpus of American beer reviews (See Section II). Our hypothesis is that the reviewers working for this website will be subject to odour-taste synaesthesia and describe flavours and aromas as such. As a consequence, certain smell descriptions in the reviews should often be accompanied by the same taste descriptions and vice versa.

This leads us to our first research question (RQ1): is it indeed the case that taste and smell descriptions are closely linked, and do beer reviewers, by consequence, use the same lexical descriptors for taste and smell?

The second research question (RQ2) is the following: are the expert beer reviews meaningful providers of information considering the limited vocabulary and ways of describing sensory perceptions such as smells, flavours, and sight? Therefore, sensory experiences should be worded in a consistent manner.In [8], the authors have shown this is the case for authors of wine reviews. They built classifiers to predict colour, grape variety, country of origin, and price of a wine, based on the experts' wine reviews. The experimental results showed promising F-scores, demonstrating that wine reviews really are informative.

In this research, we investigate whether beer experts share a common vocabulary to describe beer properties. The consistency of the descriptions will be verified by building a machine learning system to automatically predict beer properties for new beer reviews on the basis of smell/aroma and taste/flavour descriptions from experts' reviews. For these preliminary experiments, we build a system that predicts beer colour labels on the sole basis of smell and taste properties. This means that the system can assign colour properties to the beer in

 ISBN: 978-1-61208-578-4 

the review even though colour descriptions are lacking. The automatic prediction system then bases its colour property predictions on the smell and taste descriptions present in the review. The general hypothesis is then that beer experts, just like wine experts, are capable of describing beer properties in a sufficiently consistent manner, which allows beer properties to be automatically predicted on the basis of experts' reviews. Training automatic systems to predict beer characteristics could be a first step to develop content-based recommender systems for beer. Whereas current recommender systems only take metadata like beer style (e.g., IPA) and user-based filtering or subjective ratings into account, beer recommendations based on review content and aroma and taste descriptions could be very useful.

The remainder of the paper is structured as follows: Section II describes the beer review corpus we used for these experiments. Section III elaborates on the colour classification experiments for automatic prediction of colour based on smell and taste descriptions in the reviews (RQ2), while Section IV presents the lexical analysis of the language that is used for describing smell and taste in beer reviews (RQ1). Finally, in Section V, we draw conclusions and present prospects for future research.

## II. Corpus

For our experiments, a corpus of online beer reviews written by experts is composed. Experts are widely considered to be more accurate and detailed in their smell, taste and sight descriptions, which is important for the construction of an automatic prediction system. In [9], the author claims that experts are biologically superior to novices when it comes to distinguishing tastes.

We have collected 2205 beer reviews from the American website Tastings.com [10] and automatically extracted per review the following structured beer properties: *name, category, alcohol, country, style, aroma, flavor, bitterness*, as well as the review text written by the expert. Figure 1 shows an example of the structured beer properties that are listed for the different beers.

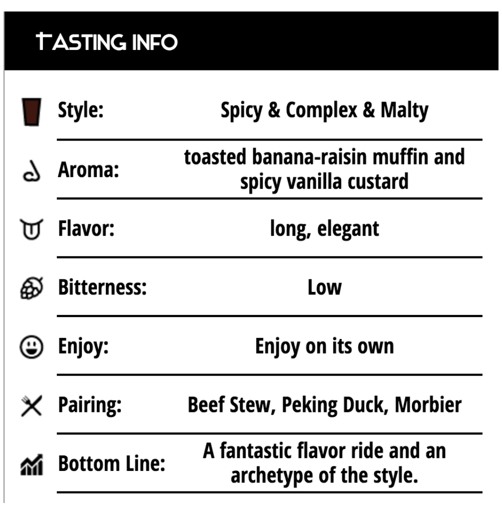


Figure 1. Example of the structured beer properties .

Example 1 lists the review text accompanying the structured properties of the *Westmalle Trappist Double* beer:

(1) *Hypnotic reddish mahogany color. Rich aromas of toasted banana-raisin muffin and spicy vanilla custard with a satiny, fruity-yet-dry medium-full body and a long, elegant finish with notes of caramelized nuts and dried fruits, peppery spice, and earth. A fantastic flavor ride and an archetype of the style.*

## III. Classification experiments

The task of predicting the colour of the beer was conceived as a supervised classification task. Two sets of bag-of-words features were extracted as information sources from the expert review texts: unigrams (single words) and bigrams (sequences of two words). As we want to investigate the viability of automatically predicting the colour of the beer based on the sole review text, sentences containing a colour description were automatically removed from the review. The reviews were further preprocessed by removing all punctuation marks and by lower-casing all words contained in the review.

Unlike the other beer properties, colour descriptions could not be automatically extracted from the website, because they were only present in the written review itself and not in the structured information about the beer in the website. Therefore, the word preceding the word *colour* in the review was extracted as the colour description. If the review contains for instance *Cloudy golden color with a high head*, the word preceding *color*, being *golden* in this case, is automatically extracted as the colour label and the entire sentence is removed from the review text. Altogether, 49 different colour labels were extracted following this procedure. This high number of classes, however, makes the colour classification of beers from new reviews less accurate. Therefore, colour terms referring to the same colour category were grouped together and 7 new classes were formed (*very light, yellow, amber, brown, black, red/rose*, and *green*). Other colour terms could not be grouped within these classes, and were kept as individual colour classes. This was the case for the colours: *deep, hazy, cloudy*, and *oak*. Some of the latter colour labels do not refer to a specific colour, but are artifacts introduced by the automatic extraction of the label. For the experiments, colour labels only occurring once in the corpus were not considered (*brilliant, violet, indigo, gray, platinum, nickel, wood*). This way we ended up with 11 different colour categories or classes, which are shown in Table I. Beer reviews where no colour label could be extracted, were removed from the database, resulting in a reduction of the corpus from 2205 to 2121 instances.

TABLE I. Colour classes.

| Colour category | Colour labels |
|---|---|
| very light | silver |
| gold | gold, yellow, golden, sunburst, straw, light, bright brassy, sunset, white, sunrise |
| amber | amber, copper, bronze, orange, maroon, penny |
| brown | brown, mahogany, medium-brown, walnut dark-brown, sienna |
| black | black, ebony, cola |
| red/rose | ruby, garnet, pink, red, salmon |
| green | green, emerald |
| cloudy | cloudy |
| deep | deep |
| hazy | hazy |
| oak | oak |

As a classification algorithm, we used Support Vector Machines as implemented in the LIBSVM toolkit [11]. As evaluation measures, we report (ten-fold cross-validated) (1)

 ISBN: 978-1-61208-578-4 

Precision, (2) Recall and (3) $F_1$-score per colour class, calculated as follows:

$$Precision = \frac{Number\ of\ correctly\ predicted\ labels}{Total\ number\ of\ predicted\ labels} \quad (1)$$

$$Recall = \frac{Number\ of\ correctly\ predicted\ labels}{Total\ number\ of\ gold\ standard\ labels} \quad (2)$$

$$F-score = \frac{2(Precision\ *\ Recall)}{Precision\ +\ Recall} \quad (3)$$

In addition, we report accuracy, which simply divides the number of true predictions (both positive and negative class) by the total number of instances of the complete data set. The scores for colour labels occurring in less than 5 training instances are not reported in the results. Due to lack of sufficient training data, these rare labels are never predicted by the classifier (and thus result in 0% performance for all evaluation measures).

### A. Experimental results

For the presented colour prediction experiments, LIBSVM was applied with two different kernels. The first experiment was run with the linear kernel of LIBSVM, resulting in an average cross-validation accuracy of 56.05%. For the second experiment, the default (RBF) kernel was optimised by means of a grid search on one training fold, resulting in an optimised $c$ parameter value of 2.0 and an optimised $g$ parameter value of 0.0078125. This second experimental setup yielded the best results, with an average accuracy of 59.88%. We compared these results to two classification baselines: (1) a majority baseline predicting the most frequent class for all instances, being *amber* and (2) a random baseline predicting labels uniformly at random. Table II shows the results of the two baselines and the two variations of our colour prediction system.

TABLE II. ACCURACY SCORES FOR TWO BASELINES AND TWO VERSIONS OF THE COLOUR PREDICTION SYSTEM.

| System | Accuracy |
|---|---|
| Baseline 1 | 39% |
| Baseline 2 | 1% |
| linear kernel | 56% |
| optimised RBF | 60% |

Table III presents the detailed results per individual colour class for the first experiment, while Table IV reports the results for the optimized classifier. As can be noticed, the F-scores for the more frequent classes (i.e., *yellow, amber, brown*) are higher in the optimised version, but the performance of *black*, which has 148 training instances, drops considerably (from 35.9% to 20.9% F-score). A second observation that can be made is that colour labels with few training instances are never predicted by the classifier, resulting in an F-score of 0%. Hence, to improve the prediction accuracy for all colour classes, more training data need to be collected to have a more balanced corpus. In future research, we will investigate alternative experimental set-ups, including an ensemble of binary classifiers trained for each colour separately.

In addition, a shallow qualitative analysis revealed that the classifier is often confused between the colour labels *amber, gold, brown* and *black*. The confusion matrix for *amber*, for instance, illustrates that the classifier often predicts amber beers as *gold* (221 times) and to a lesser extent as *brown* (75 times):

TABLE III. CLASSIFICATION SCORES REPORTING THE NUMBER OF INSTANCES (AND CORPUS DISTRIBUTION), PRECISION, RECALL AND F-SCORE ON THE POSITIVE CLASS.

| Colour category | Nr of instances (distribution) | Recall | Precision | F-score |
|---|---|---|---|---|
| very light | 9 (0,5%) | 0.0 | 0.0 | 0.0 |
| gold | 683 (32,5%) | 59.8 | 62.7 | 61.2 |
| amber | 819 (39%) | 55.7 | 62.8 | 59.0 |
| brown | 366 (17%) | 52.4 | 49.7 | 51.1 |
| black | 148 (7%) | 41.2 | 31.8 | 35.9 |
| red/rose | 41 (2%) | 0.0 | 0.0 | 0.0 |
| green | 18 (1%) | 0.0 | 0.0 | 0.0 |
| oak | 18 (1%) | 0.0 | 0.0 | 0.0 |

TABLE IV. SCORES FOR THE OPTIMIZED CLASSIFIER.

| Colour category | Nr of instances (distribution) | Recall | Precision | F-score |
|---|---|---|---|---|
| very light | 9 (0,5%) | 0.0 | 0.0 | 0.0 |
| gold | 683 (32,5%) | 64.5 | 63.8 | 64.2 |
| amber | 819 (39%) | 58.6 | 72.8 | 64.9 |
| brown | 366 (17%) | 55.5 | 57.9 | 56.7 |
| black | 148 (7%) | 55.9 | 12.8 | 20.9 |
| red/rose | 41 (2%) | 0.0 | 0.0 | 0.0 |
| green | 18 (1%) | 0.0 | 0.0 | 0.0 |
| oak | 18 (1%) | 0.0 | 0.0 | 0.0 |

TABLE V. CONFUSION TABLE OF AMBER.

| Gold standard label | Predicted label | Nr of instances |
|---|---|---|
| amber | amber | 514 |
| amber | black | 9 |
| amber | brown | 75 |
| amber | gold | 221 |

This can be explained by the fact that the different variations of these colours present a continuum, rather than a clear-cut distinction ("pale amber" resembles "deep gold", "amber brown" resembles "brown" and "deep brown" is very similar to "black" in reality). As a result, one can assume that the beer experts are not 100% consistent in naming these similar beer colours and might use similar lexical descriptors in the review text for colour variations that are alike.

### B. Most informative lexical descriptors for colour

To gain insight in which n-grams are most characterizing of colour-labelled dataset, we performed Mutual Information (MI) feature scoring. Feature selection filter metrics, such as MI, can be used to characterize both the relevance and redundancy of variables [12]. The mutual information between two random variables is a non-negative value which measures the dependency between the variables. It is equal to zero if and only if two random variables are independent, and higher values mean higher dependency. We used the MI implementation in the Scikit-learn toolkit [13] which relies on nonparametric methods based on entropy estimation from k-nearest neighbors distances as described in [14] ($k = 3$).

We used the MI-score ranking as an approximation of most characterizing features for the target colour labels in which higher MI-scores are more dependent on the 11 colour target classes. The vast majority (81.11%) of the n-grams

TABLE VI. MUTUAL INFORMATION SCORE RANKING TOP 20.

| Rank | MI score | ngram |
|---|---|---|
| 1 | 1.89E-01 | chocolate |
| 2 | 6.80E-02 | stout |
| 3 | 6.54E-02 | of chocolate |
| 4 | 5.98E-02 | coffee |
| 5 | 4.97E-02 | cider |
| 6 | 4.69E-02 | dark |
| 7 | 3.68E-02 | lemon |
| 8 | 3.58E-02 | light-to-medium body |
| 9 | 3.56E-02 | light-to-medium |
| 10 | 3.43E-02 | cherry |
| 11 | 3.28E-02 | porter |
| 12 | 3.21E-02 | dark chocolate |
| 13 | 3.11E-02 | nuts |
| 14 | 3.03E-02 | mocha |
| 15 | 2.70E-02 | apple |
| 16 | 2.47E-02 | chocolate and |
| 17 | 2.36E-02 | medium-to-full body |
| 18 | 2.35E-02 | roasted |
| 19 | 2.33E-02 | cocoa |
| 20 | 2.21E-02 | toffee |

over the 90th percentile of scores ($P_{90} = 5.85$e-3, $n = 254$) pertain unambiguously to smell or odour semantic classes. This shows that the most characteristic n-grams for the colour labels largely pertain to semantic classes of smell and odour. As illustration of this conclusion, Table VI gives an overview of the 20 highest ranked MI descriptors.

## IV. LEXICAL DESCRIPTORS FOR TASTE SMELL

In the introduction, it was hypothesized that beer reviewers are subject to odour-taste synaesthesia and by consequence use the same lexical descriptors for taste and smell. To verify this premise, a corpus analysis was performed for the lexical descriptors used in the structured "aroma" and "flavor" labels and a frequency list of all lexical descriptors assigned to both categories was compiled. The full aroma frequency ranking consists of 3121 terms, of which 1993 are unique terms (63.86% of the aroma corpus) and the full flavour ranking contains 2466 terms, of which 1456 unique terms for flavour description (59.04% of the flavour corpus). Figure 2 illustrates, however, a rapid stagnation of both lines starting from the top-500 most frequently used terms. A closer examination reveals that indeed, in the aroma ranking, 1832 terms are only used once, 389 twice and 189 three times in the entire training corpus. In the flavour ranking, 1416 terms are only used once, 310 twice and 160 three times. These are terms that are rather uncommon for describing both aroma and flavour.

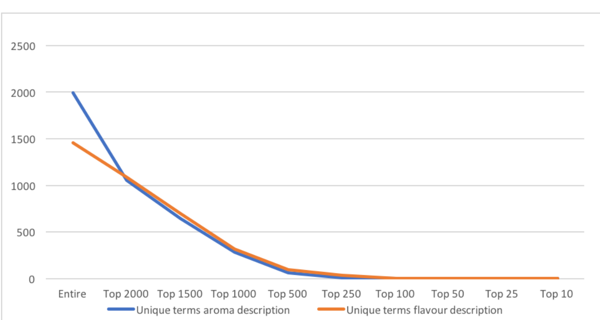


Figure 2. Distribution of unique terms for aroma and flavor labels.

Table VII lists the number of unique *aroma* and *flavour* terms (and their corresponding percentage of the respective aroma and flavour corpus). The ten most frequently used terms to describe smell properties (i.e., aroma) are also used to describe taste properties, which means that none of these terms are unique for aroma description. Only one term (*roasted nuts*) is unique in the top 25 and even top 50 most frequently used terms in the aroma ranking. In its top 100 most frequently used terms, only two terms (*roasted nuts* and *danish*) are unique for aroma description. For the description of taste properties (i.e., flavour) only one term (*tangy*) in the top 10 and top 25 most frequently used terms is unique. Two terms (*tangy* and *grassy*) are unique in the top 50 of the flavour ranking and in its top 100, five terms (*tangy*, *grassy*, *driven*, *radish sprouts* and *bitter greens*) are solely used for the description of flavour.

TABLE VII. UNIQUE TERMS FOR AROMA AND FLAVOUR DESCRIPTIONS.

| | Nr of unique aroma terms | % of corpus | Nr of unique flavour terms | % of corpus |
|---|---|---|---|---|
| top 10 | 0 | 0.000 | 1 | 0.041 |
| top 25 | 1 | 0.032 | 1 | 0.041 |
| top 50 | 1 | 0.032 | 2 | 0.081 |
| top 100 | 2 | 0.064 | 5 | 0.203 |
| top 250 | 12 | 0.384 | 35 | 1.419 |
| top 500 | 60 | 1.922 | 98 | 3.974 |

The low number of unique terms for both aroma and flavour descriptions confirm our initial hypothesis of odour-taste synaesthesia [7], which states that people recognise and describe aromas and flavours, respectively smell and taste properties, similarly. The fact that all 10 most frequently used terms for aroma description are present in the top 100 most frequently used terms for flavour description and that only two of the top 10 most frequently used terms for flavour description are unique, shows that indeed many lexical descriptors are used for both aroma and flavour descriptions.

## V. CONCLUSIONS

This paper presents preliminary research investigating the sensory descriptors used by expert beer reviewers. The performed lexical analysis confirms the odour-taste synaesthesia hypothesis, as the most frequently used descriptors are shown to be used for describing both aroma and flavour properties of beers.

In addition, we wanted to examine whether expert beer reviewers succeed at describing sensory experiences in a consistent manner. To this goal, we conducted a machine learning experiment aiming at automatically predicting beer properties, being the colour of the beer for the present research. By relying on the fact that perceptions of sight are closely related to perceptions of smell and flavour, consistency of the beer property descriptions has been shown, because the review text was the sole information source used in colour prediction. Our classification experiment showed promising results, with an average accuracy score of about 60%. Analysis of the results, however, revealed that the classifier was only successful at predicting the most frequent colour classes. Moreover, a statistical analysis by means of Mutual Information showed that the most informative review terms for colour prediction largely pertain to the semantic classes of smell and odour.

In future research, we want to collect more data in order to have a more balanced corpus and start collecting data for other languages as well. This way, we can carry out multilingual analyses of the lexical descriptors that are used to express smell and taste sensations. In addition, we will add more advanced (semantic and syntactic) features to improve the classification accuracy and perform experiments aiming at predicting other beer properties in an automated way.

## REFERENCES

[1] S. Levinson and A. Majid, “Differential ineffability and the senses,” Mind & Language, vol. 29, no. 4, 2014, pp. 407–427.

[2] I. Croijmans and A. Majid, “Not All Flavor Expertise Is Equal: The Language of Wine and Coffee Experts,” PLoS ONE, vol. 11, no. 6, 2016, e0155845. doi:10.1371/journal.pone.0155845.

[3] C. Spence and B. Piqueras-Fiszman, The Perfect Meal: The Multisensory Science of Food and Dining. Oxford, UK: John Wiley and Sons, 2014.

[4] G. Morrot, F. Brochet, and D. Dubourdieu, “The Colors of Odors,” Brain and Language, vol. 79, no. 2, 2001, pp. 309–320.

[5] M. Dematte, D. Sanabria, and C. Spence, “Cross-Modal Associations Between Odors and Colors,” Chemical Senses, vol. 31, no. 6, 2006, pp. 531–538.

[6] C. Spence, “Multisensory Flavor Perception,” Cell, vol. 161, no. 1, 2015, pp. 24–35.

[7] G. Calvert, C. Spence, and B. Stein, The Handbook of Multisensory Processes. Cambridge: MIT Press, 2004.

[8] I. Hendrickx, E. Lefever, I. Croijmans, A. Majid, and A. van den Bosch, “Very quaffable and great fun: applying nlp to wine reviews,” in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2016, pp. 306–312.

[9] L. Bartoshuk, “Comparing sensory experiences across individuals: recent psychophysical advances illuminate genetic variation in taste perception,” Chemical Senses, vol. 25, no. 4, 2000, pp. 447–460. [Online]. Available: + http://dx.doi.org/10.1093/chemse/25.4.447

[10] “Tastings.com,” 2017, URL: https://www.tastings.com/ [accessed: 2017-01-02].

[11] C.-C. Chang and C.-J. Lin, “LIBSVM: A Library for Support Vector Machines,” ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 3, 2011, pp. 27:1–27:27, ISSN: 2157-6904.

[12] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” Journal of machine learning research, vol. 3, no. Mar, 2003, pp. 1157–1182.

[13] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, 2011, pp. 2825–2830.

[14] B. C. Ross, “Mutual information between discrete and continuous data sets,” PloS one, vol. 9, no. 2, 2014, p. e87357.

# An Analysis of the Collaborative Network Mechanism Based on the Dynamic Network

Hyejin Lee, Seok-Hyoung Lee, Kwang-Young Kim
Dept. of Information Convergence Research
Korea Institute of S&T Information
Daejeon, Rep. of Korea
E-mail: hyejin@kisti.re.kr, skyi@kisti.re.kr, glorykim@kisti.re.kr

***Abstract*—This ongoing research aims to explore factors contributing to collaboration in various time intervals using the Stochastic Actor-Oriented Model (SAOM), a dynamic network analysis method. This model can be used to measure correlations between the consequences of an individual's choice and the network structure in a time series, thereby permitting simultaneous observations of a network at the micro-level and macro-level. To accomplish the research purpose, three mechanisms related to collaboration were established as hypotheses based on the literature review: reciprocity, hierarchy, and similarity. These mechanisms were combined cumulatively to determine whether they were significant for collaboration. Accordingly, the results of a pilot experiment showed that reciprocity and similarity did not have significant effects individually across time intervals, but their explanatory power about collaborative relations increased when these two variables were used in combination.**

***Keywords-Social Network Analysist; Dynamic Network; Co-author Network; Stochastic Actor-Oriented Model***

## I. INTRODUCTION

In the context of this study, collaboration is interaction between researchers and a long-standing scientific practice for advances associated with a discipline. Interaction between researchers is recognized as a part of the research process in which the research community effectively communicates and shares information about various investigations [1].

Researchers benefit from the following effects of collaboration. During the process of exchanging ideas and tacit knowledge about a research topic and reaching an agreement regarding anticipated findings and analytical methods, they can share research equipment or engage in formal/informal communications. In addition, continuous collaboration in research activities permits access to a wider range of information sources, new knowledge, and an increased pace for dissemination of research findings based on the various backgrounds of participating researchers. Further, collaboration reduces research expenses through the shared use of large-scale equipment or facilities and helps expand research activity opportunities [1] [2].

Researcher collaboration carries it own important meaning, but it also has many intrinsic values from a metrology-based data analysis perspective. Not only can researcher collaboration produce material for data analysis, but the meaning of collaboration can also be interpreted in many ways. The most frequently used metrology-based collaboration analysis method is the co-author network analysis method, which refers to social networks. It is applicable to researcher collaboration because the data can measure the scope of a collaboration easily by utilizing co-authors' unique features; additionally, it is objective, demonstrable, and stable over time [3].

Scientific findings are the product of massive collaborations; in other words, papers having multiple co-authors could be produced in massive numbers [4]. The latest research trends in "big science" and "data science" have allowed the number of collaboration-based, co-authored papers to increase explosively. For instance, the Thomson Reuters's Web of Science (WoS) indicates that—as they relate to physics—120 papers with more than 1,000 co-authors and 44 papers with more than 3,000 co-authors were published in 2011[5].

With the consistent growth in publication of co-authored papers, the co-author network continues to change dynamically, repeatedly evolving and differentiating into various forms. For this reason, a scientific network is often categorized as a dynamic network [6]. Accordingly, an analysis of the co-author network from the perspective of a dynamic network is needed [7]. The Stochastic Actor-Oriented Model (SAOM), one of the dynamic network analysis methods, is made according to processes of individual choice [8]. Also, the individual choice affects the overall collaborative network structure. This is why we adopted the SAOM in this study.

This research defines a co-author network as a dynamically changing network and tracks network changes over time based on the longitudinal analysis of a co-author network. The aim of this research is to identify empirically a relationship between individual factors and the network structure for significant factors used by individual

 ISBN: 978-1-61208-578-4

researchers to select their co-authors and establish a collaboration network.

## II. STOCHASTIC ACTOR-ORIENTED MODEL(SAOM)

To verify the research hypotheses, the SAOM—a social network analysis method—was adopted. Social network analysis focuses on a social entity—a relationship between an actor and entity—to not only identify a relationship between two individuals but also expand and explain a social relationship between individual entities. The SAOM is a computational model focused on the actor(s). It initiates a complex system based on a small virtual world consisting of many interacting actors. This virtual world is made up of the actor, the system in which the actor acts and interacts, and the external environment that affects the system. This model can be used to research the interactions between various actors and better reflect and analyze the actual world in which the characteristics or choices of an actor (an individual) could affect another individual. It is especially effective in research explaining how micro-level interaction, such as a personal relationship, affects macro-level interaction.

Research using this model in information science is still in an early stage, and no research reflecting various characteristics of the actor has been conducted to date. Kronegger et al. [9] conducted research on the community of Slovenian scientists using the SAOM, examining changes in the collaboration network structure for each time interval in four scientific fields. Ferligoj et al. [10] also applied the SAOM to Slovenian scientists in seven disciplines and explained factors changing the domestic and overseas collaboration network based on the cumulative performance expectation mechanism. Zinilli [8] used this model and investigated researcher collaboration in the Projects of National Interest (PRIN) performed in Italy. The research traced changes in network links based on four academic disciplines and used the h-index as an independent variable for choosing a partner.

## III. RESEARCH DESIGN

To achieve the research purposes, the following four steps describe the research process: 1) establish the hypotheses, 2) select data, 3) extract data, and 4) analyze and verify.

### A. Research Process

First, the hypotheses address factors for collaboration based on a literature review of existing research papers. By grouping research findings suggested by the papers, the researchers defined the top mechanisms for the formation of collaborations.

Second, data were selected to choose research subjects in the nanoscience field and set time intervals to measure networks. Subjects included both key researchers and co-authors in the field. Time intervals were set to range from 2001—the year when the National Nanotechnology Initiative was established—to 2015. To select specific fields of nanoscience and measured time intervals, researchers focused the analysis on nanoscience-related policy and used WoS and InCites databases to select data.

Third, the researchers are extracted data to generate attribute information and co-author relationship information about the research subjects. Accordingly, foundational data for data extraction came from WoS. Attribute information was based on individual characteristics derived from the hypotheses as collaboration factors, and co-author relationship information was based on the co-author relationships for published papers.

Fourth, this research incorporated a longitudinal analysis, which applied the SAOM to the extracted data and verified the hypotheses. An analysis of the overall co-author network structure and a pattern analysis of each measured time interval were performed. An analysis of factors pertaining to individual or network structure and affecting the formation of a co-author relationship was also conducted. Though NetMiner 4.0—which specializes in analyzing and visualizing overall structure and pattern—was used for the former, RSiena was used for the latter, as it specializes in longitudinal analyses of networks.

### B. Research Hypothesis

First, research collaboration establishes a network based on interrelationships and trust.

Maglaughlin and Sonnenwald [11] regard trust in researchers' research capabilities, values, and academic knowledge as important factors in choosing collaboration partners. Hara et al. [12] interpreted the reason researchers choose partners based on prior collaborative relationships as the decreased chances for failure with research outcomes. Intimacy with collaboration partners could also enhance collaboration efficiency, and the collaboration process is easier when a researcher partners with others he or she personally knows [13].

*H1*: *Researchers' academic friendship and trust (whether they co-authored in the previous year or research career) would affect the formation of an academic collaborative network.*

Second, researchers tend to forge collaborative relationships with partners who are characterized as more esteemed. As actors whose cumulative performance is expected to be high have more connections in a network structure [14], nodes are concentrated on the nodes with high levels of centrality. Abbasi et al. [15] demonstrated that degree centrality, eigenvector centrality, average strength, and network efficiency were independent variables affecting the g-index, which indicates an individual researcher's research performance.

*H2: Researcher hierarchy (organizational reputation) would affect the formation of an academic collaborative network.*

Third, the similarity mechanism is known as homophily; it means that similar characteristics and qualities between actors promote intimacy and bonds, with a high likelihood for the formation of a network structure. Various research papers have cited results for homophily in researcher collaboration. Maglaughlin and Sonnenwald [11] viewed the research topic as an important factor when researchers collaborated with other researchers. Kronegger et al. [9] stated that a network was formed depending on similarities between individual researchers' organizations.

*H3: Similarities between researchers (organizations) would affect the formation of an academic, collaborative network..*

## IV. DATA ANALYSIS AND PILOT TEST RESULTS

In order to achieve the purpose of this research, data analysis was carried out as follows and the results of the research were derived.

### A. Data Analysis Process

To extract authors and co-authors in nanobiotechnology fields as data, papers in the fields were collected by subject area through WoS and its methodology. Initially, NT papers in Korea from 2001 to 2015 were used in the first subject search scope in accordance with the WoS search rule, and the scope of subjects were reduced to BT based on the search results. As a result, the total number of collected papers was 1,704. From the collected papers, five authors who consistently published papers within the designated time interval(s) and who also had high h-indexes were selected; then, their co-authors were extracted.

Table I shows the results from the collection of papers via WoS for the five authors (applying time intervals). Data extracted included numbers of co-authors and co-author relationships. The total number of identified authors was 1,154, and the number of authors identified in the last time interval was 644.

As a next step, the researchers are extracted co-authors for Time Interval 3 (2011 – 2016) based on the publication year and established a collaborative network. Based on the same authors in Time Interval 3, collaborative networks were established in the same manner for Time Interval 1 (2001 – 2005) and Time Interval 2 (2006 – 2010). Because the population of each network should remain the same to identify objectively the factors that cause a collaborative network to look a certain way in a given time interval, the number of nodes in Time Intervals 1 to 3 were made consistent at 644; then, experimental data were generated.

TABLE I. STATISTICS OF THE COLLECTION OF PAPERS

| Number of papers | Number of authors in papers | Average number of authors per paper | Number of identified authors |
|---|---|---|---|
| 597 | 3,224 | 5.49 | 1,154 |

### B. Results of Pilot Test Analysis

#### 1) Network Structure Analysis for Each Time Interval

The network structure analysis results showed that the average degree or the average number of edges, which represents the number of collaborative relationships per researcher, increased as time passed.

Density gets closer to 1 when all researchers are linked to one another, but in this research, it decreased over time, most likely because of the rapid growth of edges.

The level of fragmentation rose gradually over time. Decreasing density means that, even though subgroups were established, their internal density was low and there were few excessive disconnects between subgroups.

Clustering represents a subgroup's level of separation, and fragmented subgroups seem to be the result of segmentation of the subject area, differentiation of the key researcher, and convergence with other fields.

TABLE II. RESULTS OF NETWORK STRUCTURE FOR EACH TIME INTERVAL

| Observance Time Interval | Degree | Mean of Degree | Density | Clustering |
|---|---|---|---|---|
| 2001-2005 | 52 | 1.529 | 0.093 | 0.781 |
| 2006-2010 | 538 | 3.183 | 0.038 | 0.84 |
| 2011-2015 | 3,591 | 5.567 | 0.017 | 0.861 |

Figure 1 shows the co-authors' network by time interval. The network structure for each time interval is composed of three components.

#### 2) Network Structure Analysis for Each Time Interval

This section longitudinally analyzes factors that influence collaborations between nanobiotechnology researchers in Korea. To examine the effect of each factor, this research will identify collaboration factors for individual independent variables as the first step and combine them phase by phase to examine the extent of increased explanatory power from each combination of variables based on changes of the random variable as the second step.

Therefore, network data 1, 2, and 3 (the N by N adjacency matrix) were used as dependent variables for each of the three intervals and dependent variables' changes in accordance with 4 independent variables in each time interval were confirmed using RSiena.

In the pilot test to analyze the collaboration mechanism involving network changes using RSiena, whether a network establishment factor was significant in the overall interval was examined. The overall maximum convergence ratio is an index for identifying whether a network establishment factor is significant in the overall interval; when the ratio is 0.25 or less, the factor is considered significant. Two out of five researchers and their co-authors were analyzed, and the results showed that the combination of two independent


ISBN: 978-1-61208-578-4

variables—“whether researchers co-authored in the previous year” and “whether researchers belonged to the same organization”—were interpreted to have a significant effect on the dependent variables, as described in the Table Ⅲ.

## V. FUTURE RESEARCH DESCRIPTION AND EXPECTED OUTCOMES

Research is now in progress to analyze significant factors regarding the five researchers’ selections of co-authors for each time interval and the structural characteristics affecting their individual choices. This research holds academic significance for its analysis of a co-author network phenomenon from a new angle based on application of the SAOM, a social network theory, to information science. More notably, the method that combined a dynamic network structure analysis and the probability theory is meaningful in expanding the traditional static network-based macro-level co-author analysis methodology and connecting the macro-level (structure) and the micro-level (actor) in a dynamic way for analysis. Moreover, the results of the analysis are worth using for service planning (e.g., information service recommendations or predictions).

## REFERENCES

[1] S. Lee, and B. Bozeman, “The impact of research collaboration on scientific productivity,” Social studies of science, vol. 35, number5, 2005, pp. 673-702.

[2] G. Abramo, C. A. D’Angelo, and F. Di Costa, “Research collaboration and productivity: is there correlation?” Higher Education, vol. 57, no.2, 2009, pp. 155-171.

[3] DH. Lee, IW. Seo, HC. Choe, and HD. Kim, “Collaboration network patterns and research performance: the case of Korean public research institutions,” Scientomerics, vol. 91, 2012, pp.925-942.

[4] G. Melin, “Pragmatism and self-organization Research collaboration on the individual level,” Research Policy, vol. 29, 2000, pp.31-40.

[5] J. Adams, “Collaborations: The rise of research networks,” Nature, vol. 490, no. 7420, 2012, pp.335-336.

[6] J.S. Katz, and B.R. Martin, “What is research collaboration?” Research Policy, vol. 26, no.1, 1997, pp.1-18.

[7] A. Perianes-Rodriguez, L. Waltman, and N. J. van Eck, “Constructing bibliometric networks: A comparison between full and fractional counting,” Journal of Informetrics, vol. 10, no. 4, 2016, pp.1178-1195.

[8] A. Zinilli, “Competitive project funding and dynamic complex networks: evidence from Projects of National Interest(PRIN),” Scientometrics, vol. 108, 2016, pp. 633-652.

[9] L. Kronegger, F. Mali, A. Ferligoj, and P. Doreian, “Collaboration structures in Slovenian scientific communities,” Scientometircs, vol. 90, 2012, pp.631-647.

[10] A. Ferligoj, L. Kronegger, F. Mali, T. A. Snijders, and P. Doreian, “Scientific collaboration dynamics in a national scientific system,” Scientometrics, vol. 104, 2015, pp.985-1012.

[11] K. L. Maglaughlin, and D.H. Sonnenwald, "Factors that impact interdisciplinary natural science research collaboration in academia," 10th International Conference of the International Society for Scientometrics and Informetrics, Jul. 2005, pp. 499-508. Stockholm.

[12] N. Hara, P. Solomon, S. L. Kim, and D. H. Sonnenwald, “An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration,” Journal of the American Society for Information science and Technology, vol. 54, no. 10, 2003, pp.952-965.

[13] D. Stokols, S. Misra, R. P. Moser, K. L. Hall, and B. K. Taylor, “The ecology of team science: Understanding contextual influences on transdisciplinary collaboration,” American Journal of Preventive Medicine, vol. 35, no. 2, 2008, pp.S96-S115.

[14] M. Ruef, H.E. Aldrich, and N.M. Carter, “The Structure of Founding Teams: Homophily, Strong Ties, and Isolation Among U.S. Entrepreneurs,” American Sociological Review, vol. 68, no. 2, 2003, pp. 195-222.

[15] A. Abbasi, J. Altmann, and L. Hossain, “Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regressionanalysis of performance measures and social network analysis measures,” Journal of Informetrics, vol. 5, no. 4, 2011, pp. 594-607.

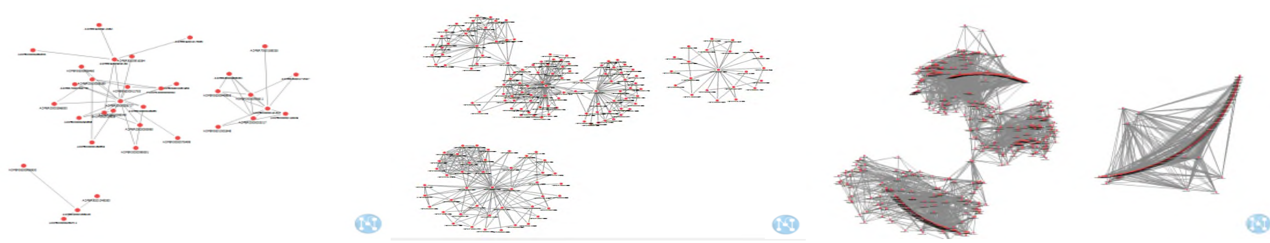

Figure 1. Results of network structure for each time interval

TABLE III. RESULTS OF SAMPLE TEST

| Variable | Measurement Method | Researcher A t-conv.max value (N=234) | Researcher B t-conv.max value (N=154) |
|---|---|---|---|
| Researcher friendship | Whether researchers co-authored in the previous year | 0.6888** | 0.1161* |
| Researcher career | Researcher’s research career in the measured time interval | 1.7564 | 1.8215 |
| Organization reputation | The organization’s contributions to the field (number of papers published) | 0.9429** | 0.9920** |
| The same organization | Whether researchers belonged to the same organization | 0.8954** | 0.1108* |

 ISBN: 978-1-61208-578-4 

| | | | |
|---|---|---|---|
| Researcher friendship + the same organization | Whether researchers co-authored in the previous year + whether researchers belonged to the same organization | 0.1953* | 0.1002* |

# Component based Agent Simulation Modeling for Self-Evolving House Market Prediction

Joonyoung Jung, Jang Won Bae, Chunhee Lee, and Euihyun Paik
Smart Data Research Group
Electronics and Telecommunications Research Institute
Deajeon, Korea
e-mail: jyjung21@etri.re.kr

***Abstract*— Because actual population forecasting is expensive and impossible, recent agent-based microsimulation is used to predict social problems. An agent-based model (ABM) models the interaction between an agent and each agent. However, long-term simulations using rule-based ABMs accumulate simulation errors. To avoid this error accumulation, the simulation model must be dynamically reconfigured using the actual data during the simulation. In this paper, we propose a component-based agent simulation modeling for model reconstruction and implemented a housing market ABM simulation system using DEVS (Discrete Event System Specification) C++ engine to evaluate the effect of error accumulation avoidance. As the simulation progresses, the housing market ABM is dynamically modified to reduce errors between the actual data and the simulation results.**

***Keywords-Agent Simulation; Component; Self-Evolvement.***

## I. INTRODUCTION

Actual population expectation is expensive and impossible because the modern society is complex and various. Therefore, microsimulation modeling (MSM) and agent based modeling (ABM) are used for modeling and simulation. Microsimulation models the individuals with real data and defines behaviors based on transition probabilities derived from micro data. ABM models individuals and interaction between the individuals. ABM mostly defines behaviors based on rule. Simulation mimics the operation of an actual process or system and is used to predict the future [1]. We can perform simulations to take appropriate action on potential problems in the future. Through micro-simulation, various characteristics of the members of society can be predicted and used to predict social problems [2] [3]. However, agent-based simulation has the disadvantage of accumulating errors, so long-term simulations can be error-prone. To avoid this error accumulation, it is necessary to reconstruct the simulation model dynamically using real data. In this paper, we propose a component-based agent simulation model for reconstructing a simulation model.

The rest of the paper is organized as follows. Section II describes related work, while Section III describes the house market simulation. In Section IV, we describe component based ABM modeling for housing market simulation, and in Section V, we provide the result of our experiments. Some concluding remarks are finally given in Section VI.

## II. RELATED WORK

MSM is one of the most useful modeling tools for simulating at micro level and for predicting the consequences of possible policy changes [2] [3]. A simulation model for predicting the future using MSM was implemented. In other words, the dynamic simulation of the income model was developed to analyze the long-term distributional consequences of retirement and aging problems in the United States [4] and the dynamic microsimulation model was developed in Australia to model economic and demographic change in the Australian population over time [5] [6]. An agent is a micro-unit that can independently determine its behavior based on environment, own state, and interaction with other agents [7]. The ABM was proposed in the segregation model [8], household demography [9] and population dynamics [10]. However, previous ABMs are not able to evolve the structure and parameters of model. Therefore, long-term simulations accumulate simulation errors.

## III. HOUSING MARKET SIMULATION

We designed and implemented a housing market ABM simulation system to evaluate the effect of error accumulation avoidance.

### *A. Self-evolving simulation framework*

The Self-evolving simulation is an agent simulation to mitigate long-term simulation error autonomously. To evolve simulation model, there are some modules in framework. The data management module (DMM) collects, digitizes, and saves data in database. The change recognition module (CRM) receives real world data and simulation results from the DMM and a simulation module (SIM). And then, it compares these data. Whenever the CRM recognizes difference over threshold, the CRM informs a model evolvement module (MEM) of the change. The MEM is triggered by the CRM when change is discovered. The MEM finds the evolving strategy to decrease the difference between real world data and simulation results. And then, it suggests that an ABM reconfiguration module (ARM) should change parameter and model with new parameter and model, which are suggested by the MEM. The ARM updates the simulation model according to the strategy from the MEM. To update a simulation model using an evolving strategy, the simulation model should be consisted of

 ISBN: 978-1-61208-578-4 

components. For simulation performance, the distributed and paralleled simulation may be accomplished by a simulation engine.

### B. *House market simulation scenario*

The self-evolution simulation framework can be applied to social problem simulations that require long-term simulations. We apply the proposed framework to a house market simulation, as shown in Figure 1. First of all, there are some agents, such as a house holder, external supplier, and bank, and some environments, such as a house market and house, in a house market ABM. The house holder is a micro agent to buy, borrow, sell, and lend a house. The external supplier is a micro agent to make, sell, and lend a house. The bank is a micro agent to give loan to the house holder who wants to buy or borrow a house. The house market is an environment to collect the house information, which is listed to buy or sell. The house is an environment to have all houses information. The real world house market data are collected. The essential data are house price and house trade volume. The economic data related to house market are gross domestic product, exchange rate, price index and money supply. Therefore, these data is received in real world and then digitized. The house policies are policy parameters, which have an effect on the house price and house trade volume. These policies are input parameters of the simulation. And then, these policies are analyzed using the simulation results.

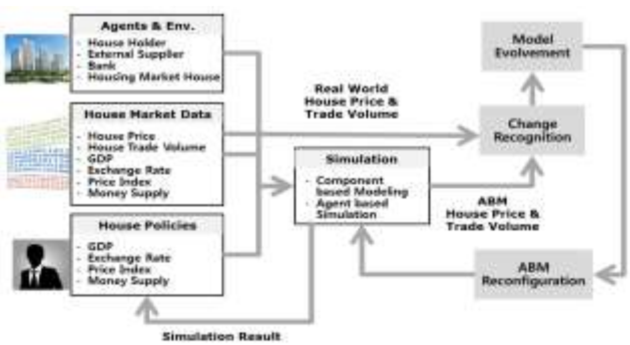


Figure 1. Hosing market simulation scenario.

The CRM receives the house price and trade volume of real world from DMM and the house price and trade volume of ABM simulation result from SIM.

CRM receives real world house prices and trade volume from the DMM, and receives the house values and trade volume of the ABM simulation results from SIM. When CRM perceives the difference over threshold between these data, CRM informs MEM of it. The MEM evolves the economic condition recognition of each agent as hidden variable. The MEM decides the hidden variable using machine learning and then sends it to ARM. The ARM updates the simulation model according to information from MEM. After updating the simulation model, the new simulation model conducts a simulation. This process is recursively repeated to reduce a long-term simulation error. After finishing the simulation, the simulation results are analyzed by policy maker or simulation analyzer.

## IV. COMPONENT BASED ABM MODELING FOR HOUSING MARKET SIMULATION

Agent model can be componentized. Each agent consists of the situation awareness (SA), adaptation and projection (AP), and decision making (DM). Each SA, AP, and DM consists of one atomic component. Components have various levels from atomic model (AM) to coupled model (CM). The model structure from the top level, housing market ABM, to the lowest level, AM, is shown in Figure 2.

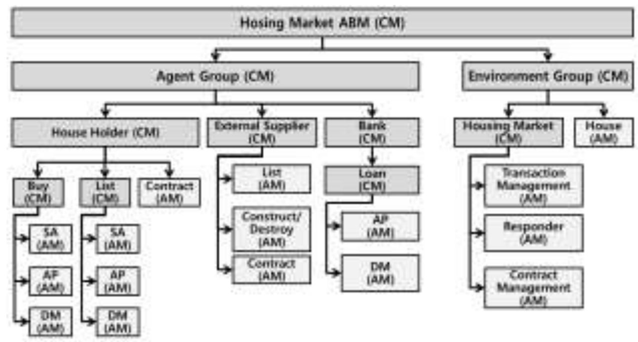


Figure 2. Housing market ABM structure.

Housing market ABM is a top level model and includes the agent group and environment group. Agent group can identify the situation and make judgment on its own as an individual subject. The agent group is divided into the house holder, external supplier, and bank. The environment group is a passive object that does not judge itself and affects agents. A house holder is a key component of this model and includes buy, list, and contract actions. For buy and list, it is designed by SA, AP, and DM concept. An external supplier has the actions of list, construct, destroy, and contract. A bank gives a loan and receives money from house holder. The loan behavior can be designed by SA and DM concept. A housing market is an environment element that connects the house holder and house. It consists of the transaction management, responder, and contract management. In these AM and CM, the essential models are the agent group and buy model.

### A. *Agent group design*

Agent group model consists of the house holder, external supplier, and bank. These agents interact with each other and environment group, as shown in Figure 3.

The house holder provides the following functions. First of all, it lists houses for sale and rental at an environment group. Second, it purchases and rents a house listed in the environment group. Third, it carries out contracts with other agent. Forth, it requests the house information listed in the environment group. Fifth, it receives information about economic indicators from the environment group. Sixth, it receives and processes contract termination information from environment group. Seventh, it receives the housing market information from the environment group. Eighth, it requests

ISBN: 978-1-61208-578-4

loanable amount to the bank and receives the response from the bank. Ninth, it requests a loan to the bank and receives the response from the bank.

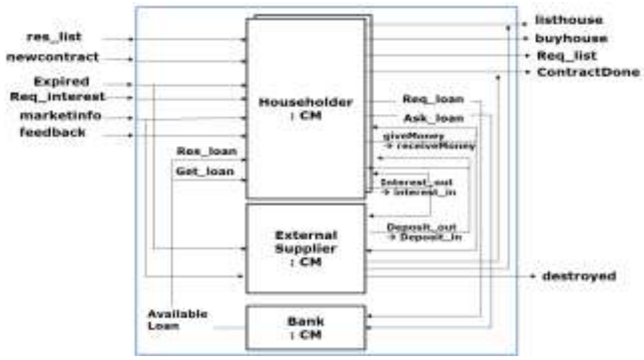


Figure 3. Agent group model for housing market ABM

The house holder provides the following functions. First of all, it lists houses for sale and rental at an environment group. Second, it purchases and rents a house listed in the environment group. Third, it carries out contracts with other agent. Forth, it requests the house information listed in the environment group. Fifth, it receives information about economic indicators from the environment group. Sixth, it receives and processes contract termination information from environment group. Seventh, it receives the housing market information from the environment group. Eighth, it requests loanable amount to the bank and receives the response from the bank. Ninth, it requests a loan to the bank and receives the response from the bank. The external supplier provides the following functions. First, it requests for demolition of the existing house listed in the environment group. Second, it builds new house and lists the new house in the environment group. Third, it receives and processes the contract termination information from the environment group. Forth, it receives the housing market information from the environment group. The bank provides the following functions. First, it carries out the function of giving loan to the agent who wants to purchase and rent the house. Second, it receives the inquiry message about loanable amount from the house holder and responds the loanable amount to the house holder. Third, it receives a message requesting a loan from the house holder and responds the loans to the house holder.

*B. Buy behavior design*

The buy behavior is consisted of SA, AP, and DM. The modeling methodology designed through SA, AP, and DM can easily reconstruct the detailed behavior of each agent. Basically, each behavior of Agent is designed as SA, AP, and DM as shown in Figure 4.

The SA can be reconfigured into any type of component if only the input and output matches the AP and DM. AP and DM can be easily reconfigured if the input and output are matched with other components.

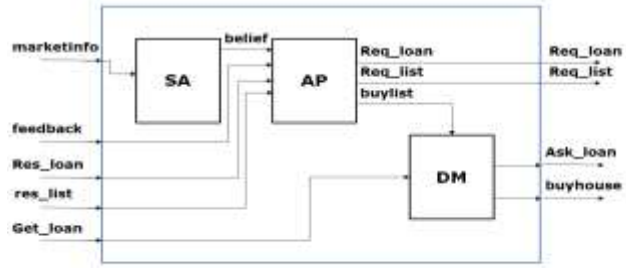


Figure 4. Buy behavior model for housing market ABM

The SA receives the market information (average transaction price, average transaction rate, etc.) of housing from the environment group and judges the real estate market situation as boom, a recession and a normal time, and then transmits the real estate market situation to AP as shown in Figure 5.

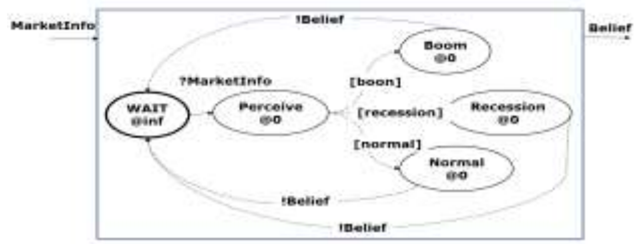


Figure 5. SA model of buy behavior for housing market ABM

The AP requests loan and house information and then receives the information on the market situation (boom, recession, normal), feedback (list of sale information, selling price, average selling price), loans, and a list of available houses. After receiving all information, it judges the house price after considering the current price and the future price, and transmits the result of judgment (buy list) to the DM as shown in Figure 6.

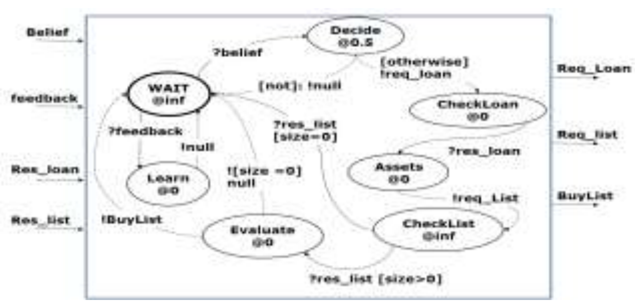


Figure 6. AP model of buy behavior for housing market ABM

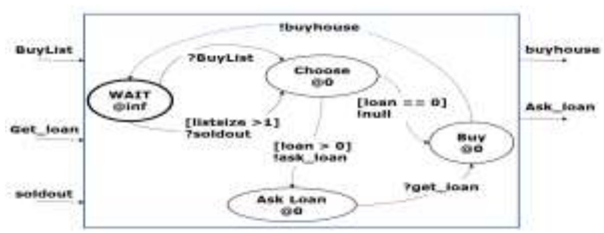


Figure 7. DM model of buy behavior for housing market ABM

The DM receives a list of houses to be purchased from the AP and requests and receives a loan amount. And then, it determines a house to be purchased using the loan amount and the house list information. The result of DM (buy house) is transmitted to the environment group as shown in Figure 7.

## V. EXPERIMENTS

In order to verify the developed housing market ABM, we implement it as a formal agent component model, and verify the reconstruction of component.

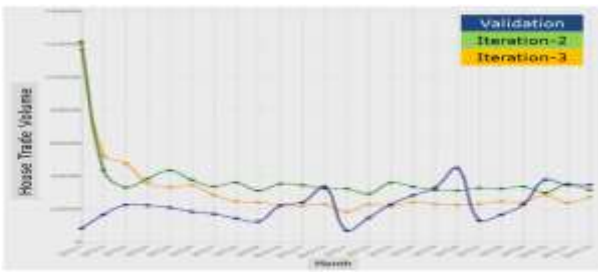


Figure 8. Model reconfiguration phase for mitigating the difference between validation data and simulation result

We implement an ABM simulation system using the discrete event system specification (DEVS) C++ engine. The initial model (iteration 1) is conducted in the simulation system and then the simulation result (house trading volume) is compared with the validation data. The MEM finds better parameter and component of model to reduce the error between validation data and simulation result and then sends result information to ARM. The ARM updates the model according to the information from MEM, and then conduct simulation recursively.

During the evolving simulation, the errors of previous model and changed model are compared as shown in Figure 8. Errors with the verification data in the iteration 2 were reduced in the iteration 3. The validation data are the real house trade volume of Seoul in Republic of Korea. We got the real house trade volume data from the web site managed by Ministry of Land, Infrastructure and Transport in Republic of Korea.

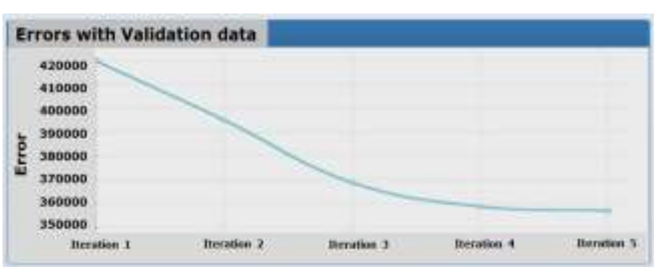


Figure 9. Self-evolving simulation result for house market ABM

After finishing the evolving iteration, the house trade volume of each model (from iteration 1 to iteration 5) is compared to see how much errors have gone down, as shown in Figure 9. As the iteration progresses form iteration 1 to iteration 5, the house trading volume errors with the validation data are reduced. Therefore, the parameter and component of model are properly modified autonomously.

## VI. CONCLUSION AND FUTURE WORK

This paper addressed the component based agent simulation modeling for self-evolving house market prediction. This paper proposed component model to reconfigure the model autonomously. To evaluate the effectiveness of the component based agent simulation modeling for self-evolving simulation, we implement house market ABM simulation system using DEVS C++ engine. The self-evolving simulation automatically updates the housing market ABM and reduces errors between the validation data and the simulation results.

We will implement the proposed simulation in the distributed and paralleled simulation environment.

## ACKNOWLEDGMENT

This research was supported by the ICT R&D program of MSIP/IITP, [R7117-16-0219, Development of Predictive Analysis Technology on Socio-Economics using Self-Evolving Agent-Based Simulation embedded with Incremental Machine Learning].

## REFERENCES

[1] J. Banks, J. Carson, B. Nelson and B. Nicol, “Discrete-event system simulation,” Prentice Hall, 1984.

[2] G. H. Orchtt, "A new type of socio-economic system," The review of economics and statistics, vol. 39, no. 2, pp. 116-123, May 1957.

[3] A. Harding, "Challenges and opportunities of dynamic microsimulation modelling," Plenary paper presented to the 1st General Conference of the International Microsimulation Association, Vienna, Aug. 2007.

[4] M. Favreault, and K. Smith, "A primer on the dynamic simulation of income model (DYNASIM3)," Urban Institute, Feb. 2004.

[5] S. Antcliff, "Introduction to DYNAMOD: a dynamic population microsimulation model," Technical Paper No 1, National Centre for Social and Economic Modelling (NATSEM), University of Canberra, 1993.

[6] A. King, M. Robinson, and H. Baekgaard, “DYNAMOD-2: An overview,” Techincal Paper No. 19, National Centre for Social and Economic Modelling (NATSEM), University of Canberra, Dec. 1999.

[7] A. Getchell, "Agent-based modeling," Physics, pp. 757-767, Jun. 2008.

[8] T. C. Schelling, “Dynamic models of segregation,” Journal of mathematical sociology, vol. 1, pp. 143-186, 1971.

[9] N. Geard, J. M. McCaw, A. Dorin, K. B. Korb, and J. McVernon, “Synthetic population dynamics: A model of household demography,” Journal of Articial Societies and Social Simulation, 16(1) 8, Jan. 2013.

[10] K. Singh, M. Sajjad, and C. W. Ahn, "Towards full scale population dynamics modelling with an agent based and micro-simulation based framework," International Conference on Advanced Communication Technology (ICACT), pp. 495-501, Jul. 2015.

# Community Works: Predicting Changes in Community Resilience

Alexandra Trani, C.J. Hutto, Dennis Folds, and Tom McDermott,
Georgia Tech Research Institute
Atlanta, Georgia
e-mail: Alexandra.Trani@gtri.gatech.edu

***Abstract*—Community resilience is a multidimensional concept that would be difficult if not impossible to measure with a single assessment. To capture this system-of-systems nature of community resilience, we argue that considerations of human and social capital must be included because humans are both the *source* of community resilience and the *beneficiaries* of it. We build on a data transformation method proposed by Hutto and colleagues [7] allowing researchers to create comprehensive measures of community resilience and its underlying social constructs (i.e., subjective well-being and objective standard of living). Using a combination of data simulation via probability sampling and confirmatory factor analysis, we demonstrate the impact of some future (conjectured) proposed legislation—e.g., governmentally provided self-driving cars as a public transportation alternative—on community resilience for three *demographically* defined communities: the elderly, the disabled, and all Americans of legal driving age (i.e., 16+) for each of the *geographically* bounded communities consisting of the 50 United States and the District of Columbia.**

***Keywords-human capital modeling, social capital modeling, prediction, human capital investments, social capital investments, disability, self-driving cars, mass transit***

## I. INTRODUCTION

Community resiliency refers to a community's ability to respond to threats and challenges, successfully adapt to changes, and prevent, mitigate, or recover from disasters [1]-[4]. Rather than being a static construct, community resiliency emerges from the harmonic interaction of the quality of existing built infrastructure, the adequacy and efficiency of community emergency response services, and the human and social capital of residents [5][6]. Despite being vital to the measurement and prediction of community resilience, human capital—that is, the knowledge, skills, and attributes of residents that provide value to a community [8]—has been frequently overlooked by community resilience researchers and is often excluded from models of community resilience [8]. For models to capture the multidimensional system-of-systems nature of community resilience, we argue that human and social capital *must* be included because communities would simply not exist without human residents.

An immediate obstacle preventing researchers from successfully integrating human and social capital into models of community resilience is the lack of a single, comprehensive measure of human and social capital that addresses all the necessary variables required to visualize and quantify the human and social side of resilience. It is unlikely that any single measure will ever be comprehensive enough to sufficiently capture human and social capital in a resiliency context [7]. Instead, we propose a data transformation technique that allows researchers and policy makers the option to transform and combine existing data from multiple sources into a single dataset that objectively assesses existing community resilience in a comprehensive and mutable way. We use this technique in the present paper to characterize *existing* community resilience as well as to *predict* the impact of future (conjectured) legislation—the availability of self-driving cars—on different communities. Our technique allows researchers to reliably create extremely representative samples of a population quickly using as many variables as necessary to address important questions about human and social capital in a community resilience context. Because our technique does not rely on any one data source for information, researchers and policy makers can update models to include the most time-relevant population statistics with relatively little effort.

We will briefly define the factors of human and social capital modeled in the present paper. Additionally, we will explain our data transformation technique so that researchers may begin immediately employing our method to improve models of community resilience. Finally, we will provide a hypothetical use-case about the impact of proposed legislation—the availability of governmentally provided self-driving cars—on different communities in the United States.

### A. Crucial Factors of Human and Social Capital

The most important first step of any community-based model is the definition of a community of interest. Communities can be defined by *geographic* identity such as geolocation or an attributed external border including town limits or people living within a common flood area. Communities can also be defined by the *demographic* attributes of their residents using personal identity (e.g., age, gender, ethnicity) and cultural identity (e.g., political and religious affiliation) to draw community boundaries.

Once a community of interest is identified, researchers and policy makers must consider the general *political* climate (i.e., the extent that residents of a community trust and feel satisfied and secure with the government, and believe civil liberties are protected), the general *economic* climate (i.e., unemployment rates, economic growth, inflation, and gross domestic product per capita), and determine the variables important for measuring human capital for the present time [3]. Human and social capital variables can be split into two

 ISBN: 978-1-61208-578-4 

main factors: subjective well-being and objective standard of living [2]. When included in the same model, subjective beliefs about quality of life can be assessed in relation to objective measures of standard of living illuminating resident biases and assessing the extent to which residents understand their present economic health [17]-[19].

Subjective well-being is a latent factor representing the beliefs, emotions, and attitudes a resident maintains in regards to their life [9-10]. Subjective well-being can be further split into four separate sub-factors [2][11][12]: *affective experiences*, *cognitive appraisals*, *global life judgments*, and domain specific *satisfaction*. Variables related to affective experiences capture a resident's trait affect regarding quality of life as well as personal factors including marital happiness. Variables related to cognitive appraisal capture resident opinions about their present socioeconomic status with respect to their ability to achieve life goals compared to other societal groups. Variables related to global life judgements capture the extent a person believes life is exciting as well as general beliefs about human nature. Finally, variables related to domain specific satisfaction capture the extent residents feel fulfilled and satisfied with life aspects such as career goals, family, and friendships. Building upon the data transformation techniques described in [7], we move beyond overly simplified measures of mood or satisfaction and assess subjective well-being as a multifaceted and complex construct.

Standard of living is a latent factor assessing a resident's *objective* access to present wealth, happiness, comfort, and material goods. At the national level, standard of living is frequently operationalized as gross domestic product (GDP) per capita. At the level of the resident or individual, standard of living can be split into two distinct sub-factors: quality of life and material wealth. Variables related to quality of life capture life expectancy, crime rates, environmental quality and living conditions, and resident's access to goods and services. Variables capturing material wealth assess resident's wages and income, net worth, cost of living, and wealth relative to neighbors. As standard of living increases, so too does subjective well-being [13]-[16].

We extend Hutto et al.'s [7] model to include *community engagement* and *social capital* as further indicators of community resilience. Community engagement is a latent factor representing the extent that residents participate in activities, groups, and relationships within their community and with broader society [20][21]. People who are more engaged within the community tend to measure higher in subjective and objective health and well-being [22]. Social capital is a latent factor representing the frequency of interactions requiring trust and cooperation that occur within a community. These interactions are done for a common, public good rather than simple personal gain [23]-[26] and are typically operationalized through volunteer and charity work (e.g., the amount donated to charity organizations) as well as acts of social trust and kindness (e.g., giving up a seat on a bus for a stranger). Resilient communities tend to be higher in social capital because it allows for a willingness to help residents in emergency contexts [25] as well as a readiness for a community to adapt to change [22].

In combination, these social community resilience factors (e.g., human capital via subjective well-being and standard of living, social capital, and community engagement) provide extensive information about the socioeconomic context within which a community exists. The ability to comprehensively measure each of these factors provides policy makers and researchers with a better understanding of not only how aspects of a resident's social and economic life impact community resilience but also the degree to which proposed changes in any given social variable will impact community health overall.

### *B. Our Community Data Simulation Technique*

Our community data simulation technique uses probability sampling to create a representative community of interest. Using this method, researchers can obtain and combine data from multiple relevant data sources to create the comprehensive and complex factors necessary to model community resilience. Upon creating a dataset consisting of both a representative sample and enough variables to adequately model the relevant community resilience factors, researchers are advised to test their assumptions using confirmatory factor analysis or structural equation modeling. Using these confirmatory methods, researchers will then be provided with factor loadings that can be used to create a series of weighted sums easily quantifying total community resilience. Once item and factor weights have been established for each variable and factor of interest, researchers can transform variables using criterion determined by suggested community changes. A new weighted sum would be created using the previously established item and factor weights as well as a combination of the unchanged variables and the transformed variables of interest. This new quantification of community resilience can be compared with previous community resilience to determine if suggested changes yielded meaningful increases in community resilience. Additionally, these results are readily transformed into visual aids to help researchers and policy makers communicate results. We will use the rest of the methods section to break down each step of this process.

For more information on suggested factor creation, we encourage researchers to read Hutto et al.'s presentation for a complex and comprehensive structural equation model of community resilience [7].

### *C. A hypothetical use-case for our data transformation technique*

To illustrate and motivate this research, we propose a hypothetical example of a government deciding whether to provide self-driving cars as an aid to national public transit services for people ages 16+. In the first scenario, self-driving cars would only be made available to people with disabilities. In the second scenario, self-driving cars would be provided to both seniors—with and without disabilities—and people with disabilities. In scenario three, self-driving cars would be made publicly available to anyone aged 16+ regardless of disability status. As such, for the purpose of this paper, the communities of interest are defined in two ways: by specific demographic attributes (i.e., age and disability status), and by geographic boundaries (i.e., the 50 U.S. states and the District of Columbia). For the present paper, we focus on how proposed changes would impact community resilience as influenced by human and social capital specifically (rather than community resilience).

## II. METHODS

We next decompose our community data simulation technique, step-by-step.

*Step 1. Identifying data sources and variables*

In the previous section, we defined the characteristics of communities of interest and proposed changes to the existing community structure. We now identify data sources containing either raw response data (i.e., the number of people who responded to a specific response option for a given variable) or probability data (i.e., the percent of people who fit within a certain socioeconomic criterion or who maintain a specific belief). To create the present dataset, we turned to polling sources including Pew and Gallup, the U.S. Census Bureau and American Fact Finder, and the General Social Survey.

During this step, researchers may be tempted to exclude variables that aren't directly relevant to proposed changes. We caution against this behavior. The only way that factors of interest can meet our criterion of being comprehensive, multidimensional, and representative is if they include the variables that adequately capture the construct of interest. Not all items may be related to proposed changes, however, all items should adequately, *uniquely*, and comprehensively represent their underlying factor. In this way, researchers can be more comfortable in assuming factors were not created in a way to bias findings to support or refute proposed changes.

*Step 2. Create a dataset using probability sampling*

Using the probabilities obtained in step 1, we next create a dataset of any size and with any number of potential variables by simulating data using probability sampling. Here, we select a variable (e.g., disability status), determine the categories of interest (e.g., disabled, not disabled), and use existing (verified) population percentages to set the probabilities that a person of a specific sociocultural criterion (e.g., race, sex, age, income, etc.) would be in a specific category. Using statistical software such as R's "sample function", this type of sampling can be performed with near infinite repetitions to generate a population sample of substantial size and extremely representative variability.

*Step 3. Determining Factor Loadings for items and sub-factors of interest*

With an adequate simulated community dataset in place, we next obtain item and factor loadings via confirmatory factor analysis or structural equation modeling. This is accomplished for all data items relevant to an underlying sub-factor, and for all sub-factors related to a given higher-order factor. To complete this, categorical variables will need to be transformed into continuous variables using the method described by Hutto and colleagues [7]. The relationship between items and factors should be determined using theory—we used a proposed model of community resilience to guide our item/factor relationships [7].

*Step 4. Create weighted sums of sub-factors*

Because factor loadings represent the degree of association between items and their underlying factors, factor loadings represent an item or sub-factor's weight in a weighted sum or factor score. While factor loadings will not be identical across datasets, similar communities should yield similar factor loadings if samples were created from accurate and representative statistics—especially if created using the same probabilities. Using factor loadings as item weights, we next create a series of weighted sums that, when combined, represent a quantification of community resilience—or human and social capital in the present paper. For example, we can now create a weighted sum using the factor loadings for items representing material wealth and quality of life separately. Using factor loadings for the material wealth and quality of life sub-factors, a weighted sum of standard of living can be obtained. Using factor loadings for standard of living and subjective well-being, a weighted sum of human and social capital can be obtained that will accurately represent underlying human and social capital for any given community. Factor loadings are obtained by performing a series of confirmatory factor analyses as described in step 3.

*Step 5. Create new variables representing changes in human and social capital*

Once item and factor weights are established, relevant variables can now be recoded per the criterion determined by our hypothetically proposed self-driving car legislation. For example, the availability of governmentally provided self-driving cars in rural areas may allow disabled peoples without access to public transportation services the ability to obtain reliable and affordable transportation, and thus, seek and obtain job opportunities and healthcare services previously unobtainable due to distance. Thus, researchers can assume unemployment rates and commute times among disabled peoples may decrease while general health, income, and social group participation among disabled peoples may increase simply because adequate transportation has been made available.

At this point, researchers may become concerned that transformations may over-exaggerate the impact of proposed changes in legislation. This is a valid fear. We suggest two methods to address this fear. First, researchers should use existing research as well as sound logic to guide assumptions *made a-priori* about the variables selected for transformation and change criterion while avoiding the temptation to change all variables or to alter change criterion after the fact. This helps to ensure that only relevant variables are transformed to be both reliable and objective. Second, researchers should

 ISBN: 978-1-61208-578-4 

test for the impact of multiple scenarios with different change criterion established *a-priori* to account for small, medium, and large effects. For example, the present paper accounts for small changes in commute times (commute times for people taking public transit reduced by ten minutes), medium changes in commute times (commute times for people taking public transit reduced by 20 minutes) and large changes in commute times (commute times for people taking public transit reduced by 30 minutes) to create better models for how communities may change assuming different goals are met.

*Step 6. Test that changes in human and social capital are significant and then model the changes*

We next apply statistical techniques to test the extent that proposed changes have a significant and meaningful impact on community resilience—or, in our case, human and social capital. This method allows results to be readily represented by visual aids for easy communication to any audience. For communities defined in terms of geographic boundaries, we use choropleth maps to visualize changes in data values using easy-to-see differences in color across a region of interest.

### A. *Present study*

For each state in the United States and the District of Columbia, using a deterministic draw of 1000 females and 1000 males, we simulate an age, ethnicity, and disability representative community. Extending the model developed by Hutto and colleagues [7], we next incorporate more than 100 variables thought to represent community resilience, including variables related to human and social capital. These variables are drawn from myriad data sources as previously described, and constrained to the years 2000-2016 (for the General Social Survey and Census data), or the most recent published public dataset available.

We hypothesized that impacts of government-subsidized self-driving cars on social aspects of community resilience may be *small*, *average*, or *widespread*. For demonstration purposes we apply deterministic "what-if" modeling and simulation using single-point estimates – based on a combination of inductive reasoning and empirically informed heuristics – for each level of impact. *Small* impacts were operationalized in our model by a single unit of increase for directly relevant variables – that is, for example, people may move from "Very Dissatisfied" to merely "Dissatisfied" with their present commute times. In the *average* impact scenario, variables directly related to transit (e.g., commute time/satisfaction and group attendance) would increase by two units while other, more indirect measures that may increase because of readily available and easily accessible transit (e.g., the confidence that a person could find a job as good as the one they presently have) would increase by a single unit. In the *widespread* impact condition, variables directly related to transportation were increased by three units (or until a person hit maximum satisfaction), variables that improve as an indirect result of the availability of new forms of mass transit were increased by two units, and general measures of satisfaction were increased by a single unit.

To investigate the effect on broader community resilience via impacts to human and social capital, these simulated impacts were applied to each scenario of interest—e.g., providing government-subsidized transportation using self-driving cars to either 1) the disabled community alone, 2) the combined communities of disabled and elderly, or 3) to all Americans age 16 and above. If, when compared to no change in mass transit availability, the impact was significantly different, then proposed legislation would be considered "effective".

## III. RESULTS AND LIMITATIONS

In the first scenario, we wanted to see the impact of governmentally provided self-driving cars on disabled people in America. To accomplish this, we selected variables related to the social aspects of community resilience that were likely to be influenced when disabled people are suddenly able to travel greater distances in areas where public transportation or cost effective methods of transit were not previously available. Fig. 1 demonstrates predicted changes in community resilience for each projection. Specifically, we show projected changes in community resilience for three populations (disabled, disabled and older adults, all drivers) in three different scenarios (small, average, and widespread change). Maps demonstrate linear transformations as projected community resilience increases within population for each scenario. All nine projections significantly improved community resilience from baseline predicted values.

Governmentally provided self-driving cars improved community resilience in disabled populations compared to the baseline predicted value of 107.21 (S.D.=3.04). Small scenarios improved community resilience by 17.03 points ($t(50)=40.5$, $p < 0.05$). Average scenarios improved community resilience by 51.03 points ($t(50)=121.34$, $p < 0.05$). Widespread scenarios improved community resilience by 91.03 points ($t(50)=216.45$, $p < 0.05$).

In the second scenario, we wanted to see the impact of governmentally provided cars on both disabled as well as elderly adult populations across America. There was a significant improvement in community resilience compared to baseline. Small scenarios improved community resilience by 14.76 points ($t(50)=34.47$, $p < 0.05$). Average scenarios improved community resilience by 40.76 points ($t(50)=95.19$, $p < 0.05$). Widespread scenarios improved community resilience by 75.76 points ($t(50)=176.93$, $p < 0.05$).

In the third scenario, we wanted to predict the change in community resilience associated with all Americans of legal driving age (i.e., 16+) having access to self-driving cars as a mass-transit alternative. The availability of this form of mass transportation had sweeping positive impacts on community resilience even in the small change conditions when compared to baseline. Small scenarios improved community resilience by 18 points ($t(50)=42.35$, $p < 0.05$). Average scenarios improved community resilience by 51 points

 ISBN: 978-1-61208-578-4 

($t$(50)=119.98, $p < 0.05$). Widespread scenarios improved community resilience by 92 points ($t$(50)=216.43, $p < 0.05$).

We did not weigh the sample size drawn from each state, thus, sparsely populated states like Alaska had the same number of people in our sample as densely populated states such as California because we did not weight the number of people selected from each state based on state population. Despite this, our sample was created to be as representative as possible and was rather large (2000 people per state) so we do not believe that weighting our sample size would influence our findings beyond the point of relevancy.

This method presents a viable series of predictions of what could happen when public policy aimed at improving community resilience is implemented into a given community. The next logical step in our research is to attempt to validate our prediction models and improve our data simulations and transformations using real-world data. For this to happen, researchers must coordinate with policy makers to predict the impact of proposed changes before they are implemented and, once implemented, compare prediction to real-world changes in community resilience. This is a crucial step in advancing community resilience research and it is one that, to our knowledge, has yet to be attempted. In the interim, we encourage researchers to set multiple a-priori criterion levels (e.g., small, average, and widespread impact predictions) to account for a range of possible outcomes. We caution researchers against including only those variables in a prediction model that would change based on proposed policy, and instead, argue for including a wide range of variables that best represent their underlying factor. We also emphasize that decisions to alter variables should be based on experience and supporting evidence in the literature. Adopting this approach should improve predictions and help to keep objectivity while the present method is tested.

## IV. Conclusion

Community resilience is a multidimensional process that emerges from the successful interaction of multiple systems including built infrastructure, community services, and human and social capital. We present a data transformation technique that allows researchers to combine information from multiple data sources into a single, comprehensive measure of community resilience (or any system from which community resilience arises).

Our data transformation method involves simulating data through probability sampling, creating weighted sums, and testing for change values. Using three hypothetical scenarios, we demonstrate the efficacy of our method by demonstrating the impact of governmentally provided self-driving cars on human and social capital within communities across America. Results can be readily transformed into visual aids including choropleth maps.

Our method addresses many concerns about the sparsity of survey data and the lack of a cohesive measure of community resilience and its underlying systems using perfectly representative population samples with extremely large Ns. Additionally, our method is inexpensive and expedient, allowing for researchers and policy makers to quickly test the impact of proposed legislation on communities of interest without the administration of a separate measure or survey. Finally, our method is mutable allowing models of community resilience to rapidly change in response to changes in socioeconomic climates. Together, we believe this method represents the first step in next generation prediction of community resilience and its underlying systems.

## References

[1] R.S. Burt, *Structural holes: The social structure of competition*. Harvard university press, 2009.

[2] D. J. Folds, “A notional model of creation, maintenance, and enrichment of human capital,” Presented at the 6th International Conference on Applied Human Factors and Ergonomics (AHFE), 2015.

[3] D.J. Folds and V.M. Thompson, "Engineering human capital: A system of systems modeling approach." *System of Systems Engineering (SoSE), 2013 8th International Conference on*. IEEE, 2013.

[4] G. Windle, "What is resilience? A review and concept analysis." *Reviews in Clinical Gerontology* 21.02 (2011): 152-169.

[5] T. McDermott et al., "A human-focused, holistic model of community resilience." *INSIGHT* 19.4 (2016): 66-69.

[6] G. S. Becker, Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)* (pp. 15-28). The University of Chicago Press, 1994

[7] C.J. Hutto, A. Trani, E. Williams, H.S. Shim, D.J. Folds, T. Mcdermott, and M. Nadolski, “The social side of community resilience”. *HUSO*, 2017.

[8] J. Mincer, "The distribution of labor incomes: a survey with special reference to the human capital approach." *Journal of Economic Literature* 8.1 (1970): 1-26.

[9] F.M. Andrews and J. P. Robinson, "Measures of subjective well-being." *Measures of personality and social psychological attitudes* 1 (1991): 61-114.

[10] E. Diener, S. Oishi, and R. E. Lucas, "Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life." *Annual review of psychology* 54.1 (2003): 403-425.

[11] E. Diener, "Assessing subjective well-being: Progress and opportunities." *Social indicators research* 31.2 (1994): 103-157.

[12] E. Diener, S. Oishi, and R. E. Lucas. "Subjective well-being: The science of happiness and life satisfaction." *Oxford handbook of positive psychology* 2 (2009): 187-194.

[13] M. Blaxter, *Health and lifestyles*. Routledge, 2003.

[14] D.G. Blazer et al, "The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey." (1994).

[15] S. Weich, G. Lewis, and S. P. Jenkins, "Income inequality and the prevalence of common mental disorders in Britain." *The British Journal of Psychiatry* 178.3 (2001): 222-227.

[16] D. P. Goldberg, and P. Huxley, *Common mental disorders: a bio-social model*. Tavistock/Routledge, 1992.

[17] E. Angner, "The politics of happiness: subjective vs. economic measures as measures of social well-being." (2008).

[18] E. Diener and M.EP Seligman, "Very happy people." *Psychological science* 13.1 (2002): 81-84.

[19] D. Kahneman et al, "Toward national well-being accounts." *The American Economic Review* 94.2 (2004): 429-434.
[20] B. A. Pescosolido, J. D. McLeod, and W. R. Avison. "Through the looking glass: the fortunes of the sociology of mental health." *Mental health, social mirror*. Springer US, 2007. 3-32.
[21] T. R. Prohaska, L. A. Anderson, and R. H. Binstock. *Public health for an aging society*. JHU Press, 2012.
[22] W.N. Adger, "Social capital, collective action, and adaptation to climate change." *Economic geography* 79.4 (2003): 387-404.
[23] J. S. Coleman, (1994). *Foundations of social theory*. Belknap Press of Harvard University Press.
[23] R. R. Dynes, “Social Capital: Dealing With Community Emergencies,” *Homeland Security Affairs*, *2*(3), 2006
[24] Y. Nakagawa, and R. Shaw, “Social Capital: A Missing Link to Disaster Recovery,” *International Journal of Mass Emergencies and Disasters*, *22*(1), 5–34, 2004
[25] B. Wellman and S. Wortley, “Different Strokes from Different Folks: Community Ties and Social Support.” *American Journal of Sociology*, *96*(3), 558–588, 1990
[26] C.J. Hutto, “Blending quantitative, qualitative, geospatial, and temporal data: Progressing towards the next generation of human social analytics”. *HUSO*, 2016.

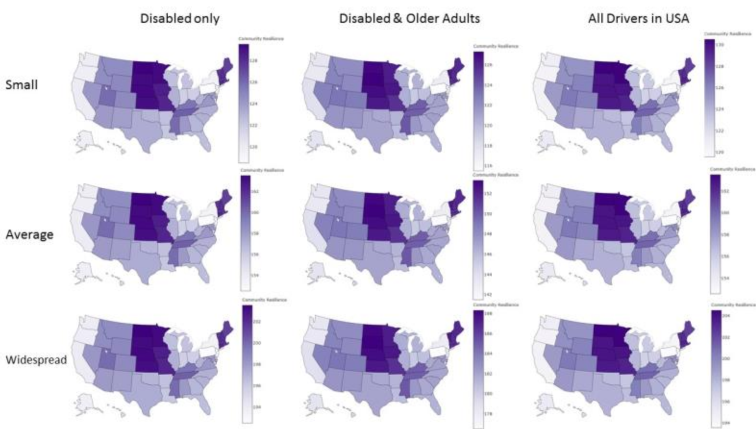


Figure 1. Predicted changes in community resilience:

# The Social Side of Community Resilience: Human Capital Modeling

C.J. Hutto, Alexandra Trani, Elizabeth Williams, Hi Shin Shim, Tom McDermott,
Dennis Folds, and Molly Nadolski
Georgia Tech Research Institute
Atlanta, Georgia
e-mail: cjhutto@gatech.edu

***Abstract*—Present measures of community resilience – that is, how communities respond or adapt to changes as well as recover from disasters – are often too shallow and fail to account for the gamut of variables contributing to community health. We argue that this problem stems from attempting to measure community resilience with an overly simplistic assessment. It is understandably difficult to construct a predictive model of community resilience. Such a model would need to be composed of variables that represent a range of elements which capture the community's ability to respond to and/or overcome natural or man-made disasters/disruptions, including factors spanning the resilience or (in)vulnerability of houses and buildings, roads and bridges, emergency services, electrical grids, computer and information exchange networks, potable water distribution systems, sanitation systems, and so on. Furthermore, the resilience associated with the aggregate human/social spirit of a community is often marginalized or, in some cases, ignored completely. The disparate nature of such a broad range of variables is that they are measured on different scales, with incongruent units, collected from diverse sources, at dissimilar time intervals. The current paper addresses all three of the challenges associated with (1) incorporating human and social elements of community resilience, (2) representing the complexity of community (social) resilience variables in a single *common latent variable construct* model that addresses concerns about disparate scales, units, sources, and types of data, and (3) creating useful models for both *characterizing* and *predicting* the resilience of a given community. We achieve this by demonstrating a novel technique for translating extant data such that the entire gamut of relevant variables are expressed in terms of their impact on human capital. Our technique then utilizes structural equation modeling techniques to construct causal (and thus, descriptive and predictive) models of community resilience.**

***Keywords-human capital modeling, social capital modeling, structural equation modeling, transformation techniques***

## I. INTRODUCTION

Global changes including international tensions and climate related stresses increasingly impact American communities. Policy makers and the public are also increasingly concerned with the health of American cities in regards to aging American infrastructure amidst rapid technological developments. To address these problems, policy makers and researchers have begun investing in community resilience—that is, investments towards understanding and improving how communities respond and adapt to changes as well as recover from disasters [1]–[3]. Improvements in physical engineering and city infrastructure are often the first consideration for improving community resilience. For instance, improved road systems allow for easier access both in to and out of a city, earthquake-proof buildings reduce risks of structure collapse in the event of an earthquake, and investments in emergency services increase community response time in the event of disasters or threats. Engineered infrastructures in urban communities are necessary ingredients of community resilience in the presence of stressors such as such as economic downturns, natural or man-made disasters, or Carrington Event-like phenomena. However, the ultimate criteria for resilience are the preservation (or restoration) of the human population affected by such stressors. Indeed, it is for the benefit of the human population that infrastructure systems even exist; it is human welfare and quality of life which are ultimately served by fortifying critical infrastructures against stressors.

Social science has a well-established body of literature demonstrating the strong relationship between an individual's *social resilience* and the role of protective factors related to their assemblage of health, well-being, and livelihood 'competencies', 'assets', 'resources', or 'endowments' collectively referred to as *human capital* [1][2][4]–[9] and *social capital* [10]–[13]. Thus, human and social capital are both recognized as crucial in achieving resilience and, through their dynamic interplay, enable a community to respond positively to risks and alter or reduce the effects of adversity [3][6]. Furthermore, individuals are responsible for maintaining and increasing human capital within communities [5][14], and human capital is vital for economic growth [15]. As human capital increases, community conditions improve and new opportunities become available for individuals [16], [17]. In turn, individuals who are higher in human capital are more likely to recognize and exploit new opportunities when they become available within the community [18].

As such, we argue that humans are both the *source* of community resilience and the *beneficiaries* of it. Nevertheless, a major, often unconsidered, aspect of community resilience is social and human capital. A paucity of adequately comprehensive measures of human and social capital creates a major obstacle in assessing community resilience. Present measures fail to capture the intricate relationships between objective quality of life measures, subjective well-being, general political and economic climates, and community demographic factors. Indeed, traditional systems-of-systems models integrating aspects of engineered infrastructures with human behavior are often over-simplified representations of what in actuality are very complex aspects of the social and physical world [19].

 ISBN: 978-1-61208-578-4 

A single measure or survey of human and social capital is an impractical goal for researchers; it is unlikely any one measure could be both comprehensive enough and time-efficient to administer. We embrace more multifaceted representations of human behavior with more complex models. Our model of human capital assimilates data of disparate forms, using disparate units of measure, collected from disparate sources, at disparate scales, and integrates them for the purpose of developing a complex, system-of-systems representation of community health, well-being, and livelihood. Importantly, our complex models are (first and foremost) explicitly motivated by extant scientific literature, and further derived based on data-driven insights from well-established public data sets comprising records from 30 different collection activities spanning 42 years (from 1972 to 2014) across nine different divisions of the United States Census Bureau and assimilated into a single data repository called the General Social Survey (GSS) [20], as well as historical data from the U.S. Bureau of Labor Statistics [21][22] and the World Bank Open Data repository [23].

Using this well-pedigreed data model (c.f., Section II), in Section III we then present a technique for transforming the multifaceted, disparate data into ordinal measures of a *single common construct*: human capital. Once transformed, we next employ advanced statistical techniques in Section IV to characterize both the strength and direction of relationships of community resilience factors (i.e., human and social capital, economic climate, political climate, etc.), which allows us to capture causal (thus, descriptive *and* predictive) models of the social side of community resilience. Section IV presents the results of our initial SEM analysis and discusses some of the limitations associated with the approach. Section V concludes by situating our work in current and prior literature, and makes recommendations on future directions for this sort of research.

## II. Social Factors of Community Resilience

We define social aspects of community resilience using a combination of theory and available data (e.g., survey data, public reports, and scientific findings). To begin, we briefly elaborate on the theoretical definitions for the community resilience factors in the present study. Also, we elaborate upon the types of data employed to create these factors.

### A. Constituants of Human Capital: SWB and SOL

Folds and Thompson [2] argue that human capital is a complex latent (i.e., not directly observable) construct thatcan be split into two factors: objective quality of life measurements (i.e., *standard of living*) and *subjective well-being* measures. These two factors can be further broken down into sub-factors to account for the broad range of resilience-protective factors related to the assemblage of health, well-being, and livelihood for a given community:

*1) Subjective well-being (SWB):* The subjective emotions and attitudes a person maintains in regards to their own life are collectively referred to as “subjective well-being” [24][25]. Using the GSS and other public data resources, we integrate at least 25 manifest indicators of general happiness and overall satisfaction with their personal life. The manifest indicators are organized into latent variable constructs representing four principal constituents of subjective well-being [26][27], as initially operationalized by Folds and Thompson [2] for use in our human capital modeling efforts:

a) *Affective Experiences*: the longer-term experiences of pleasant affect (as well as a lack of unpleasant affect) as indicated, for example, via a person’s general perceived happiness in life, in their marriage, and with their cohabitation companion (e.g., partner or roommates).

b) *Global Life Judgements*: a person’s judgements about their sense of purpose and general feelings of optimism towards the future. Examples of global life judgements include a person’s overall belief regarding how interesting they find their own life in general (e.g., whether they consider life to be dull, routine, or exciting), and judgements about the general nature of humanity (whether they believe most other people to be trustworthy, fair, and helpful).

c) *Cognitive Appraisals*: a person’s subjective self-assessment of their own current socioeconomic state relative to their life goals, as well as broader social comparisons. Determinants include financial status self-appraisals, appraisals regarding their career and wages, social status self-appraisals (e.g., social rank and social class), and self-appraisals regarding the relative quality of their domicile.

d) *Domain Specific Satisfaction*: the degree of fulfillment or contentment with important social elements such as satisfaction with their family life, friendships, recreational interests, job, health, and their city of residence.

*2) Objective measures of quality of life and standard of living (SOL):* Measures of subjective well-being should be complemented with objective measures like income and property value when evaluating community health and livelihood [28]–[30]. When used in tandum with subjective measures such as SWB, objective measures of standard of living (SOL) allow researchers to assess the degree to which a person’s beliefs about present life conditions (e.g., a person’s belief that they are in the upper-middle class) maps on to objective information about their tangible present life conditions (e.g., actual earning wage and property value). We operationalize SOL using 17 indicators from the GSS data to capture objective measures of individual quality of life and standard of living in our human capital model. These indicators include, for example, records of individual’s highest education level attained, the number of people living in their household, type of dwelling (and whether owned or rented), various employment characteristics (part time, full time, student/homemaker, unemployed, retired, etc.), and constant (i.e., annual inflation adjusted) income in dollars. Standard of living is closely connected with subjective well-being—that is, decreases in objective standard of living results in reduced subjective feelings of well-being and increased mental health risks which, which then in turn can further reduce objective standards of living [31]–[35].

In addition to SWB and SOL, we must also consider other important elements of the social side of community resilience. The next section addresses many of these additional factors.

### B. Other Socially Oriented Community Resilience Factors

Together, subjective well-being (SWB) and standard of living (SOL) capture the respective subjective and objective aspects of Human Capital; but, this is only a part of what comprises the social side of community resilience. We must not neglect consideration of individual and community demographics, nor the greater context the community; we need to account for demographic information as well as social perceptions of the national economic conditions, the general political climate, and the general security atmosphere.

1) *Demographics:* these characteristics form the basis by which "communities" are defined in the first place. As such, it is important to have access to information at the individual and aggregate level about community demographics including aspects of *personal identity* (e.g., ethnicity, age, gender, marital status), *geographic identity* (e.g., city/community size and geolocation), and *cultural identity* (e.g., political and religious affiliations, preferences and practices). Using items from the GSS, we operationalize community demographics at the national and regional levels.

2) *Larger contextual environment:* the general *political* climate incorporates public opinion regarding the utility and morality of national programs (e.g., satisfaction with healthcare and transportation services) as well as general attitudes about the government (e.g., public trust and perceptions that social liberties are protected). The general political climate influences—and is also influenced by—the general *economic* climate and the general *security* climate. To represent the general economic climate, we access national level historical information from the U.S. Bureau of Labor Statistics [21][22] and the World Bank Open Data repository [23]. This includes annually recorded economic data such as national unemployment rates, consumer price indices, inflation rates, prime lending interest rates, and annual gross domestic product (GDP) per capita. Our representation of the general security climate incorporates community exposure to crimes, feelings of fear, and beliefs of the efficacy of the court system as constructed from data immediately available within the GSS.

## III. METHODS

### A. Principal Data Source

The General Social Survey has been administered to a representative sample of the American public from 1972 through 2016. The present study used data from the years 1972 through 2014. Survey items include feelings about national spending, community safety, membership and engagement in social groups, income and subjective feelings of financial health. We selected this survey because it offers a long-term examination of social changes within American communities while also providing us with important information that can be adapted to represent the theoretical underpinnings of human and social capital. The GSS consists of hundreds of questions with varying degrees of hypothesized relationships to community resilience. For instance, a person's astrological sign is unlikely to be indicative of their subjective well-being or objective quality of life. For this reason, we went through each question surveyed on the General Social Survey looking for those that were most representative of our theoretical conceptions of the community resilience factors. Items were selected based on evidence from existing scientific literature and a-priori hypotheses about the theoretical makeup of identified community resilience factors.

### B. Transforming Dissimilar Data Into a Common Form

Responses to selected items on the GSS were next transformed into ordinal variables—that is, we translated (typically categorical) data into their corresponding linearly ranked associated to either SWB or SOL. Transformations in this case were generally informed by the GSS data itself (e.g., people of lower income score lower in objective quality of life than people with higher income; people with more positive global life judgements score higher on subjective well-being than people with more neutral or negative global life judgements). All ordinal transformations were further vetted using a top-down approach where we identified predominant scientific studies examining the relationship between variables of interest and their hypothesized community resilience factors. Using age variables as an example, we ordinally transformed age response categories in terms of their hypothesized human and social capital clusters, resulting in 5 groups with people between the ages of 15 and 20 having the lowest value (i.e., "1") and people between the ages of 60 and 75—retired and still healthy—having the highest value (i.e., "5").

This transformation step has three important characteristics: (1) it relies on a *systematic*, *principled*, and *scientifically-grounded* mapping of survey item responses to their appropriately ranked (ordinal) impacts on a common construct (human capital), (2) it is *extensible* to any data type, as long as a relationship can be defined in terms of direction and magnitude of influence on a factor or sub-factor within the human capital model (3) once transformed, it allows researchers to employ advanced statistical techniques (such as structural equation modeling) to create causal models with *predictive* capabilities. Fig. 1 illustrates the social aspects of community resilience; the full list of factors considered for the model (discussed in Section II) consists of 68 variables. Because documentation for these factors is voluminous (it is more than 26 pages alone), we provide the factors, the GSS survey items (and response options) associated with those factors, and literature and data-derived rationale for the transformation of the typically non-linear (categorical, or nominal) data into linear (polytomous, or ordinal) data for our model in a supplementary package accompanying this paper.

 ISBN: 978-1-61208-578-4

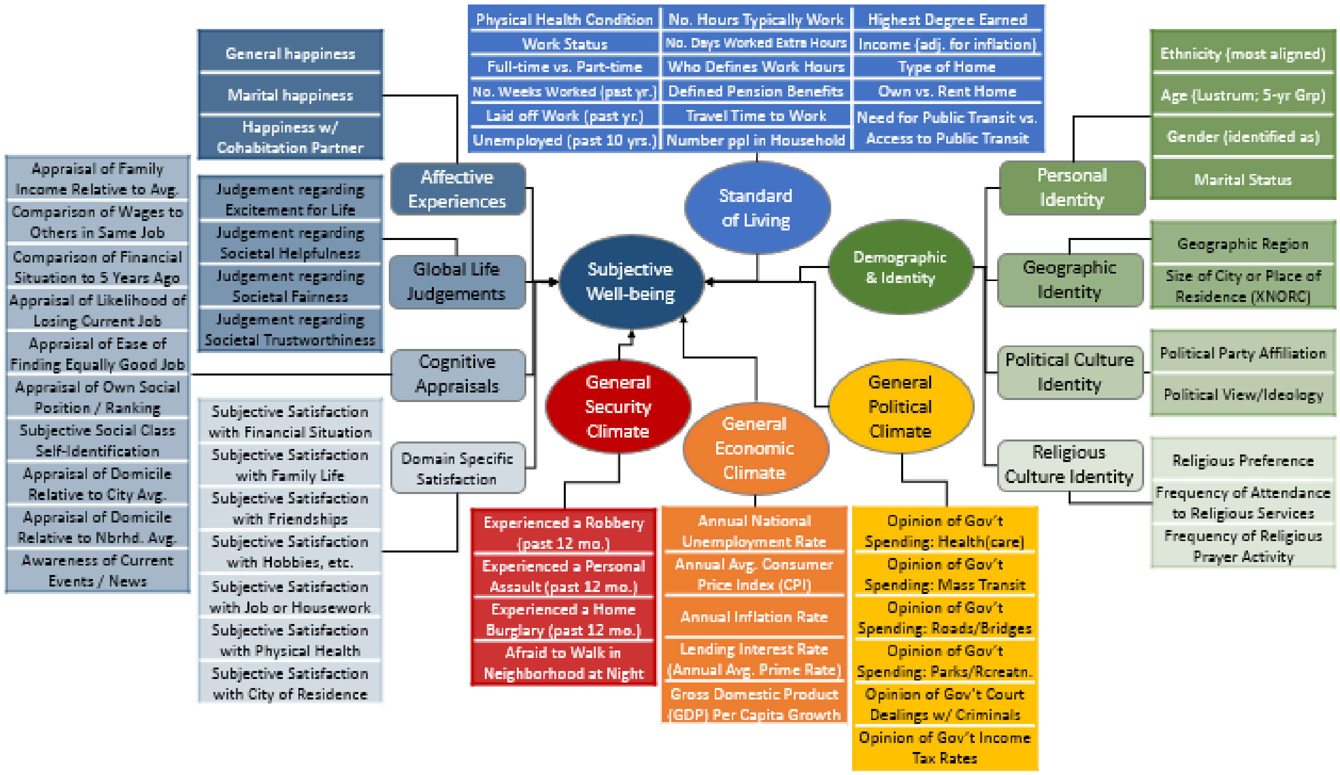


Figure 1. Factors associated with the social side of community resilience.

*C. Causal/Predictive Modeling for Community Resilience*

Any ordinally transformed variable—like the age variable in our earlier example—can be employed in complex statistical analyses including structural equation models of community resilience. Structural equation modeling (SEM) allows for objective item selection based on the degree to which items "load" onto their respective factors (i.e., the strongest predictive relationships to underlying theoretical factors will have the highest factor loadings). Additionally, SEM models provide information about the viability of hypothesized relationships between variables and factors via model fit indices. All our statistical analyses were performed in R to make our findings easily accessible, replicable, and repeatable.

SEM provides information about the direction and strength of relationships between variables and factors (either directly observable or latent) while assessing the viability of causal relationships among variables and factors [36]. Relationships are assumed to be linear so that changes at the start of a path result in linear changes in variables or factors at the end of a path [37]. To accomplish this, structural equations are computed allowing relationships to be both tested and graphically modeled [38]. When modeling causal relationships or in situations where unknown amounts of error exists in variables and factors of interest, SEM is generally superior to regression [38] making SEM a popular and accepted technique for behavioral and social modeling[39][40]. Because SEM uses a confirmatory approach toward hypothesis and model testing [37][38], it has proven to be an ideal method for modeling hypothesized human capital [14].

The best fitting SEM model is also the most parsimonious model because it accounts for the most variance between factors using the fewest causal paths. Generally, when sample sizes are large—as would be expected in human and social capital contexts—the $\chi 2$ test is biased and, while still reported, is generally not used to assess model fit [37][38], [41]–[46]. Instead, we rely on the comparative fit index (CFI; [47]) and root mean square error of approximation (RMSEA; [48]) to assess model fit. The CFI tests complete covariation between a hypothesized model and actual data providing a value constrained between 0 and 1.00. Values greater than 0.95 generally indicate a well-fitting model [38][42]. RMSEA—the best measure of model fit [38]—examines the extent that models fit a hypothesized population covariance matrix [49]. Discrepancies between population and model covariance matrices are reported as a number constrained between 0.00 and 1.00. Models with RMSEA values between 0.05 and 0.08 are considered to have adequate fit; models with values less than 0.05 have nearly ideal model fit [49].

SEM also provides information about the viability of hypothesized relationships between ordinally transformed survey items and community resilience factors. That is, the items that best represent a hypothesized factor will also have the highest factor loading in the model. Items with weak or non-existing (i.e., not significant) relationships to hypothesized sub-factors can be objectively eliminated using this approach.

Ordinal transformation also allows for aggregate examinations into community resilience factors of interest. Once transformed, survey items can be furthered centered to the mean (i.e., subtracting a mean constant from a variable of interest) and then objectively combined (i.e., summed) to form an indicator of total community resilience in an area. Total scores for individual community resilience factors can also be examined allowing for easy graphical representation (e.g., what would be accomplished using a choropleth map) for factors of interest in communities or regions of interest.

## IV. RESULTS

We transformed items from the General Social Survey from years 1972 through 2014 into polytomous, ordinal measures of human and social capital, as well as the general political economic, and security climates. We tested the relationships between these variables and their hypothesized factors using a structural equation model (SEM). Model fit statistics regarding the $\chi 2$ test for goodness of fit and the root mean square error (RMSE) for the model were adequate, though the low score for the comparison fit index (CFI) indicates that a model derived from less sparse data—e.g., a model that incorporates additional data sources along with the GSS—would likely be a better fit ($\chi^2 = 251851$, $df = 1476$, $p < 0.00$; CFI = 0.445; RMSEA = 0.053). We explore such models in subsequent research [50]. Fig. 2 shows a graphical representation of the latent variables within our structural equation model, as well as information about the strength of our hypothesized relationships between factors (e.g., subjective well-being, standard of living, demographic data, and the general political, economic, and security climates – all factors are statistically significant).

As is common when relying on a single data source, questions on the General Social Survey were not consistently asked across years, leaving information for many important items unavailable. Thus, even comprehensive surveys like the GSS are often not enough to model community resilience when used in isolation. We argue that future researchers would greatly benefit from either (1) combining multiple surveys, reports, and data sources to create a fully comprehensive measure or, (2) simulating data based on available population statistics. The methods presented in this paper support either initiative; we explore the latter in subsequent research [50].

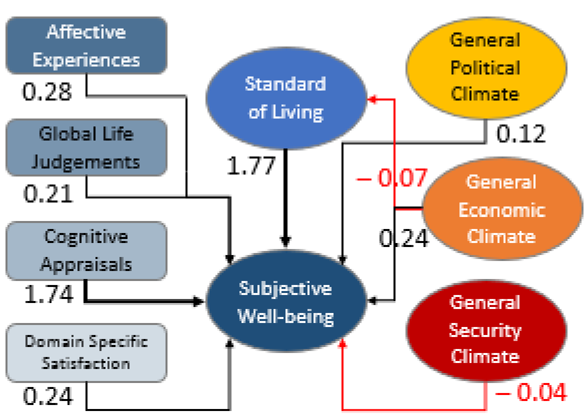


Figure 2. SEM latent variable model with factor weightings.

## V. CONCLUSION

We present a technique for transforming disparate survey data into measures of human and social capital in a community resiliency context. This technique allows researchers flexibility to create complex and representationally accurate models of human and social capital using readily available data. Our technique fulfills many of the requirements for advancing social science research, including methods for enabling researchers to analyze data consisting of huge sample collected over multiple points in time (i.e., large-N and multiple-T; [19]). We also advance social science research because our technique can be quickly utilized to extend exploratory and predictive analyses for researchers interested in human and social capital [19]. Researchers interested in applying this technique for data exploration and prediction should refer to a subsequently submitted paper [50].

Our technique was tested using Folds and Thompson's [2] structural equation model of human and social capital. We also incorporated model structure proposed by McDermott and colleagues [51] who argue that community resilience is composed of an interaction of systems including human and social capital, built environments, and city infrastructure. Our technique and model fit statistics demonstrate reasonably good support for these existing models of human capital and community resilience.

We argue that researchers using this technique in the future—especially those researchers using our technique for simulated data—should incorporate a structural equation model to both check findings and provide further tests of reliability and replicability.

## REFERENCES

[1] D. J. Folds, "A notional model of creation, maintenance, and enrichment of human capital," presented at the 6th International Conference on Applied Human Factors and Ergonomics (AHFE), 2015.

[2] D. J. Folds and V. M. Thompson, "Engineering Human Capital: A System of Systems Modeling Approach," presented at the 8th International Conference on System of Systems Engineering, 2013.

[3] G. Windle, "What is resilience? A review and concept analysis," *Rev. Clin. Gerontol.*, vol. 21, no. 02, pp. 152–169, May 2011.

[4] M. Bartley, I. Schoon, R. Mitchell, and D. Blane, "Resilience as an Asset for Healthy Development," in *Health Assets in a Global Context: Theory, Methods, Action*, A. Morgan, M. Davies, and E. Ziglio, Eds. New York, NY: Springer New York, 2010, pp. 101–115.

[5] G. S. Becker, *Human capital: a theoretical and empirical analysis, with special reference to education*, 3rd ed. Chicago: The University of Chicago Press, 1994.
[6] J. G. Dyer and T. M. McGuinness, "Resilience: Analysis of the concept," *Arch. Psychiatr. Nurs.*, vol. 10, no. 5, pp. 276–282, Oct. 1996.
[7] A. S. Masten, K. M. Best, and N. Garmezy, "Resilience and development: Contributions from the study of children who overcome adversity," *Dev. Psychopathol.*, vol. 2, no. 04, pp. 425–444, Oct. 1990.
[8] J. Mincer, "The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach," *J. Econ. Lit.*, vol. 8, no. 1, pp. 1–26, 1970.
[9] T. W. Schultz, "Investment in Human Capital," *Am. Econ. Rev.*, vol. 51, no. 1, pp. 1–17, 1961.
[10] J. S. Coleman, "Social Capital in the Creation of Human Capital," *Am. J. Sociol.*, vol. 94, pp. S95–S120, 1988.
[11] R. R. Dynes, "Social Capital: Dealing With Community Emergencies," *Homel. Secur. Aff.*, vol. 2, no. 3, Jul. 2006.
[12] Y. Nakagawa and R. Shaw, "Social Capital: A Missing Link to Disaster Recovery," *Int. J. Mass Emergencies Disasters*, vol. 22, no. 1, pp. 5–34, 2004.
[13] B. Wellman and S. Wortley, "Different Strokes from Different Folks: Community Ties and Social Support," *Am. J. Sociol.*, vol. 96, no. 3, pp. 558–588, 1990.
[14] M. Ruzzier, B. Antoncic, R. D. Hisrich, and M. Konecnik, "Human capital and SME internationalization: a structural equation modeling study," *Can. J. Adm. Sci. Rev. Can. Sci. Adm.*, vol. 24, no. 1, pp. 15–29, Mar. 2007.
[15] M. Novak and S. Bojnec, "Human Capital and Economic Growth by Municipalities in Slovenia," *Manag. Glob. Transit.*, vol. 3, no. 2, pp. 157–177, 2005.
[16] N. Bontis and J. Fitz-Enz, "Intellectual capital ROI: a causal map of human capital antecedents and consequents," *J. Intellect. Cap.*, vol. 3, no. 3, pp. 223–247, 2002.
[17] J. S. Coleman, *Foundations of social theory*. Belknap Press of Harvard University Press, 1994.
[18] P. Davidsson and B. Honig, "The role of social and human capital among nascent entrepreneurs," *J. Bus. Ventur.*, vol. 18, no. 3, pp. 301–331, May 2003.
[19] C. J. Hutto, "Blending Quantitative, Qualitative, Geospatial, and Temporal Data: Progressing Towards the Next Generation of Human Social Analytics," in *Proceedings of the Second International Conference on Human and Social Analytics (HUSO)*, Barcelona, Spain, 2016.
[20] T. W. Smith, P. V. Marsden, M. Hout, and J. Kim, "General Social Surveys, 1972-2014 [machine-readable data file]." NORC at the University of Chicago [producer and distributor], 2014.
[21] The U.S. Bureau of Labor Statistics, "Consumer Price Index Data from 1913 to 2016," *U.S. Inflation Calculator*, 2016. [Online]. Available: http://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/.
[22] The U.S. Bureau of Labor Statistics, "Labor Force Statistics including the National Unemployment Rate (Current Population Survey - CPS)," *BLS.gov Databases, Tables, and Calculators*, 2016. [Online]. Available: https://www.bls.gov/data/.
[23] World Bank, "GDP per capita (current US$) and Lending interest rate (%) from 1972 to 1016 [Custom cross-tabulation of data]," *World databank: World Development Indicators*, 2016. [Online]. Available: http://databank.worldbank.org/data/reports.aspx?source=2&series=NY.GDP.PCAP.CD&country=USA.
[24] F. M. Andrews and J. P. Robinson, "Measures of subjective well-being," *Meas. Personal. Soc. Psychol. Attitudes*, vol. 1, pp. 61–114, 1991.
[25] E. Diener, S. Oishi, and R. E. Lucas, "Personality, Culture, and Subjective Well-Being: Emotional and Cognitive Evaluations of Life," *Annu. Rev. Psychol.*, vol. 54, no. 1, pp. 403–425, Feb. 2003.
[26] E. Diener, "Assessing subjective well-being: Progress and opportunities," *Soc. Indic. Res.*, vol. 31, no. 2, pp. 103–157, Feb. 1994.
[27] E. Diener, E. M. Suh, R. E. Lucas, and H. L. Smith, "Subjective well-being: Three decades of progress.," *Psychol. Bull.*, vol. 125, no. 2, pp. 276–302, 1999.
[28] E. Angner, "The Politics of Happiness: Subjective vs. Economic Measures as Measures of Social Well-Being," *SSRN Electron. J.*, 2008.
[29] E. Diener and E. P. Seligman, "Very Happy People," *Psychol. Sci.*, vol. 13, no. 1, pp. 81–84, 2002.
[30] D. Kahneman, A. B. Krueger, D. Schkade, N. Schwarz, and A. Stone, "Toward National Well-Being Accounts," *Am. Econ. Rev.*, vol. 94, no. 2, pp. 429–434, Apr. 2004.
[31] M. Blaxter, *Health and lifestyles*. New York, NY: Routledge, 2003.
[32] D. G. Blazer, R. C. Kessler, K. A. McGonagle, and M. S. Swartz, "The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey," *Am. J. Psychiatry*, vol. 151, no. 7, pp. 979–986, 1994.
[33] D. P. Goldberg and P. Huxley, *Common mental disorders: a bio-social model*. London ; New York: Tavistock/Routledge, 1992.
[34] H. Meltzer, B. Gill, M. Pettigrew, and K. Hinds, "OPCS surveys of psychiatric morbidity in Great Britain," 1, 1995.
[35] S. Weich and G. Lewis, "Material standard of living, social class, and the prevalence of the common mental disorders in Great Britain," *J. Epidemiol. Community Health*, vol. 52, no. 1, pp. 8–14, Jan. 1998.
[36] P. M. Bentler, *EQS structural equations program manual*. Encino, CA: Multivariate Software, 1995.
[37] J. C. Loehlin, *Latent variable models: an introduction to factor, path, and structural equation analysis*, 4th ed. Mahwah, N.J: L. Erlbaum Associates, 2004.
[38] B. M. Byrne, *Structural equation modeling with EQS: Basic concepts, applications, and programming*, 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2006.
[39] R. P. Bagozzi and Y. Yi, "Multitrait–multimethod matrices in consumer research: Critique and new developments," *J. Consum. Psychol.*, vol. 2, no. 2, pp. 143–170, Jan. 1993.
[40] R. C. MacCallum and J. T. Austin, "Applications of Structural Equation Modeling in Psychological Research," *Annu. Rev. Psychol.*, vol. 51, no. 1, pp. 201–226, Feb. 2000.
[41] D. W. Gerbing and J. C. Anderson, "Monte Carlo Evaluations of Goodness of Fit Indices for Structural Equation Models," *Sociol. Methods Res.*, vol. 21, no. 2, pp. 132–160, Nov. 1992.
[42] L.-T. Hu and P. M. Bentler, "Evaluating model fit.," in *Structural equation modeling: Concepts , issues, and applications*, R. H. Hoyle, Ed. Thousand Oaks, CA: Sage, 1995, pp. 76–99.
[43] K. G. Jöreskog and D. Sörbom, *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International, 1993.
[44] R. C. MacCallum, M. W. Browne, and H. M. Sugawara, "Power analysis and determination of sample size for covariance structure modeling.," *Psychol. Methods*, vol. 1, no. 2, pp. 130–149, 1996.
[45] H. W. Marsh, J. R. Balla, and R. P. McDonald, "Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size.," *Psychol. Bull.*, vol. 103, pp. 391–410, 1988.
[46] J. S. Tanaka, "Multifaceted conceptions of fit in structural equation models," in *Testing structural equation models*, K. A. Bollen and J. S. Long, Eds. Newbury Park, CA: Sage, 1993, pp. 10–39.
[47] P. M. Bentler, "Comparative fit indexes in structural models.," *Psychol. Bull.*, vol. 107, no. 2, pp. 238–246, 1990.
[48] J. H. Steiger and J. C. Lind, "Statistically based tests for the number of common factors," in *Annual meeting of the Psychometric Society*, Iowa City, IA, 1980, vol. 758, pp. 424–453.
[49] M. W. Browne and R. Cudeck, "Alternative ways of assessing model fit," in *Testing structural equation models*, K. A. Bollen and J. S. Long, Eds. Newbury Park, CA: Sage, 1993, pp. 445–455.
[50] A. Trani, C. J. Hutto, D. J. Folds, and T. A. McDermott, "Community Works: Predicting Changes in Community Resilience," in *Proceedings of the Third International Conference on Human and Social Analytics (HUSO) (submitted)*, Nice, France, 2017.
[51] T. McDermott, D. Folds, C. J. Hutto, and M. Nadolski, "A Human-Focused, Holistic Model of Community Resilience," *INSIGHT*, vol. 19, no. 4, pp. 66–69, Dec. 2016.