

THE APPROPRIATENESS OF THE FACTUAL DENSITY AS AN INFORMATIVENESS MEASURE FOR ONLINE NEWS

Filipo Sharevski

Department of Computer and Information Technology,
Purdue University, West Lafayette, USA
f.sharevski@purdue.edu

ABSTRACT

In circumstances where the receptivity of the online news is affected by the media bias in covering public attention events, the quality of the textual component is of pervasive importance for a reliable perception of their informativeness. Aware of this threat, several natural language processing techniques have been developed for the purpose of capturing the quality of the web content based on the concepts of objectivity classification and stylometric features, knowledge maturing, factual density, or simple word count. This paper explores the appropriateness of the factual density as an adequate quality measure of the information reported on the missing Malaysia Airlines Flight 370 as a public attention event. The results suggest that the factual density needs to be applied under strict conditions in terms of increased confidence level of the textual news content, if its substance is a subject of capitalization as a referent source of information.

KEYWORDS

Factual Density, Natural Language Processing, Text Informativeness

1. INTRODUCTION

Being today's most attractive content delivery platform [1], the global Internet undoubtedly is the central location where "almost all consumers search for information in order to better inform their decisions" whether for making a purchase, read the news, realize their travel plans, improve the personal health or manage their personal finances [2]. Trumping the other content and information delivery platforms or channels when it comes to influence" [2] suggest that the information retrieved from any Internet source have a significant impact on the way our world views and perceptions are shaped with an intensive consumption of the online content [3]. Thus, the information retrieval in the digital realm needs to rely not just on the plausibility of the content itself, but should also take into consideration the "credibility and reliability of the source and the quality aspects characteristic for this content" [4], [5]. Given that the online information is characterized with a high receptivity for a narrower and subjective purpose, the impact of the information produced and disseminated through the online news on the formation of the general public brings an interesting topic for discussion.

In this context, of practical interest is to assess the quality of online media coverage of a certain event of a great public attention, together with the reliability and the plausibility of the respective information for the purpose of creating a reference article (i.e. Wikipedia article). However, the quality assessment of the online media coverage in a whole is “too general and practically impossible” [5], given the various modalities included in the production of the news, such as text, video, audio, pictures or different animations. Therefore, it is useful to restrict the online news quality assessment only to the textual portion of the news’ web feeds, since the text is the skeleton bearing the critical information about given event, supported by the additional set of pictorials or videos (in fact, many related news videos render the same content included in the textual part only in an audio-visual format). As such, the focus of the undertaken effort rests on the idea that “the quality of a text document can be related to its informativeness, i.e. the amount of useful information contained in a document” [5].

Unlike the other metrics designed to capture the quality of the web content based on objectivity classification and stylometric features [6], knowledge maturing [7], or simple word count [8], the concept of informativeness incorporates “complex semantic features” in order to better instrument the textual quality assessment [4]. According to [4] and [5], the “informativeness of a document can be measured through factual density of a document, i.e. the number of facts contained in a document, normalized by its length”. The rationale for using the factual density comes from the fact that it assesses the influential aspects of the textual news portion in more general terms compared to the objectivity classification, and that the online news reception precedes the transformation process of “contextualized information artefacts into explicitly linked and formalized learning objects”, as suggested by [7]. Moreover, it has been proven that it outperforms the word count measure [4], [5] in terms of web content quality assessment.

With this in place, the online news’ informativeness assessment is directed towards the estimation of the actual appropriateness of the factual density as an adequate quality measure of the information reported about an event of great public attention. Given the unprecedented public attention drawn by “the largest-ever multinational air-sea search” [9] for the missing Malaysia the information relative to this incident is set at the centre of the web news’ informativeness analysis. For this purpose, the online news reports collected from two different sources – Fox News and Thompson Reuters – and compared to the factual density each of the information reporting trails disseminates in the public realm of Internet journalism. Each of the datasets includes a daily review article respective to the Malaysia Airlines Flight 370 within the period of 33 days after the aircraft disappeared from the regular route, elaborating on the progress of the multinational air-sea search by bringing additional facts and potential evidence leads. The factual density of each of the respective articles in the two data sets – automatically calculated using the ReVerb [10] – is compared relative to the overall ranking of each of the articles in two instances: one where the threshold for confidence scores of the automatic factual determination is not taken into consideration, and another where this confidence threshold is considered in context of achieving an optimal balance between the recall and the precision of the extracted factual relations. The outcome of the two comparisons is finally discussed respective to the appropriateness of the use of the factual density as a measurement instrument for assessing the quality of the online news reports.

2. FACTUAL DENSITY AS AN INFORMATIVENESS MEASURE

The factual density $f_d(t)$ of an arbitrary textual resource t is the ratio between the fact count $f_c(t)$ and the size of the text $size(t)$. As pointed in [4] and [5], the most applicative approach for operationalization of the factual density is the utilization of the “Open IE methods, justified by the fact that the extraction output brings the “relational tuples representing facts from text [subject of the extraction] without requiring a pre-specified vocabulary or manually tagged training

corpora” [5], which in turn “makes such systems scalable to a broader set of relations and to a far larger corpora as the Web” [11]. Assuming a model of textual resource “as a bag-of-relations occurring in all facts extracted from it”, the relational tuple as a unit output of the Open IE systems is structured as a “triplet consisting of two arguments in the form of noun phrases and a relational phrase expressed as a verb phrase” [4], i.e. $f = (10 \text{ nations, failed to turn up, any trace of the Boeing 777 - 200R})$. Having the unit dividend in the factual density formula, the actual value of the fact count can be derived by simple summation operation of all relational tuples produced as an outcome of the Open IE system.

Since the basic definition brings no explicit details on the actual granularity of the size variable, there are three possibilities for text quantization: the overall character count including the white spaces, the word count, and the number of the sentences in the text. The most appropriate text quantization considers the word count as the text size, given the fact that it mainly strikes balance between potential miscalculation of the factual density due to an excessive normalization when the character count is used as a quantifying unit, or potential factual density degradation due to a considerably coarse quantization step in case of a sentence count. In case where the factual density is calculated relative to the character count, a potential threat exists in terms of text manipulation as to decrease (or potentially) increase the number of characters, yielding to an overall text degradation both in semantic and syntactic terms. On the other side, using the sentence unit quantization may mimic an actually large factual density of texts that are semantically and syntactically rich, but contain very few sentences. Having the word count already proposed as an actual quality measurement for web contents [8], the rationale for its choice is also valid considering a potential cross-metrics comparison.

In regards to the general threat of factual density degradation, it is also useful to reconsider the Open IE systems employed for factual derivation. In fact, “the output of the Open IE systems contains a large portion of uninformative and incoherent extractions” [5], [11] that are anyway determined as a relational tuple, which are also considered as facts that add into the final fact count and lead to overestimation of the factual density in a given textual resource. Incoherent extraction refers to a relation tuple or phrase that “has no meaningful interpretation”, coming as a result from the fact that “these systems often yield to an incomprehensible predictions” [10]. On the other side, the uninformative extraction refers to an extraction that omits critical information and is caused by “improper handling of relation tuples that are expressed by a combination of a verb with a noun, such as the light verb constructions” [10]. As a countermeasure to these threads, [10] introduced two syntactic constraints: (1) “every multi-word relation phrase must begin with a verb, end with a preposition, and be a contiguous sequence of words in the sentence”; and (2) “a binary relation phrase ought to appear with at least a minimal number of distinct argument pairs in a large corpus”, which essentially enable factual derivation “without requiring a pre-specified vocabulary or manually tagged training corpora” from a given text resource. Incorporating these improvements intended to minimize the overestimation of the factual density, Authors in [10] developed the ReVerb extractor that takes as input “noun-phrase chunked sentences from an arbitrary text resource and produces relational tuples representing the facts included within it” [12]. Given that the ReVerb “outperforms the other Open IE systems [13], [10], by achieving considerably higher precision and recall” [5][10][12], it is chosen as tool for the automatic factual density derivation in executing the central analytical engagement.

3. RELATED WORK

The factual density as a informativeness measurement of a given textual web content was initially employed by [4] and [5] in their “performance comparisons of the factual density against the word count measurement in respect to identification of featured/good articles in Wikipedia”, and the usability evaluation “of relational the features for categorizing Wikipedia articles into

featured/good versus non-featured ones”. The outcome of the performance comparison provided with an evidence that the factual density “corroborates the good performance of word count in Wikipedia in case the article are with variable length”, but in the case of articles of similar lengths “the word count measure fails while factual density can separate between them with an F-measure of 90.4%”.

Trying to assess whether the aforementioned conclusions are applicable beyond the specific class of online textual content such as Wikipedia articles, [5] “extend the analytical effort into the sphere of the arbitrary Internet texts” retrieved from various, randomly selected online contents by comparing a human-ranked reference corpus against an analogous raking derived from an Open IE factual density estimation system, consisting of 50 arbitrary web text on the material in Spanish. Correlating the “ground truth human-annotated ranking with the ranking gathered from the automatic prediction”, [5] were able to confirm the adequacy of the factual density as a measurement of the informativeness, at least in the case of Spanish web texts.

In contrast to these efforts, this work approaches the web content informativeness measurement - based on the factual density - as to assess the “actual feasibility of the automatic factual density estimation” [5] for text documents containing online news reports, instead of Wikipedia articles or arbitrary web content. By shifting the research commitment beyond the specific domains analysed by [4] and [5] towards a segment of an English-based web content of considerable interest especially in the case of public opinion formation [2], the undertaken analytical effort takes a recent event with massive public attention and news coverage - the incident involving the missing Malaysia Airlines Flight 370 and the ongoing largest-ever multinational air-sea search - as to determine whether an automatic evaluation of the informativeness of different online news feeds brings plausible and unified reporting information that can be utilized for the purpose of deriving a reliable incident briefing or creating a reference article.

4. EXPERIMENTAL SETUP AND RESULTS

Two separate datasets corresponding to the daily reporting on the investigation and search progress were compiled from the Fox News and Thompson Reuters. The Fox News feed is taken as referent since it was continuously ranked 1st in the news ranking for more than 147 months, including March, 2014 [14], which corresponds with more than 72 % of the time period of online news coverage of the incident that was initially reported on March 8, 2014. The addition of Thompson Reuters reflects the idea of balancing the potential thread of polarized online coverage [15]. Each of the datasets includes a daily review article about the Malaysia Airlines Flight 370 in a period of 33 days (between March 8 to April 9, 2014) after the aircraft disappeared, elaborating on the progress of the multinational air-sea search. Special attention was devoted to collect those daily articles that are covering the relevant topic in the overall progress in that particular point of time. The choice for this timeframe rests on the fact that the flight data recorder (able to record up to 25 hours of flight information) contains a so-called “underwater locator beacon” that can emit beacon ping signals (on 37.5 KHz) up to 30 days from the moment of the incident [16], assuming maximum operational lifetime of the flight data recorder battery (the additional reports from April 7th to April 9th as account for possible fluctuations in both the recorder’s battery lifetime and the underwater locator beacon signal).

After the textual content of all the 66 web articles were manually extracted (as to ensure that the substance of the overall report is preserved), the factual density of each of these articles in the two data sets was automatically calculated. In the first instance, the threshold for confidence scores is not taken into consideration for the factual extraction in order to determine the general degree of online news informativeness. Since ReVerb also outputs a confidence score for each of the factual extractions, in the second instance a confidence threshold of 0.6 is considered in the context of

achieving an optimal balance between the recall and the precision of the extracted factual relations [17] as an improved measurement of the informativeness. Another reason for considering a considerable higher confidence threshold value is that “it reduces the need for human fact checking” [17], which corresponds with the fact that the informativeness assessment and the comparison outlined in this work does not involve human-annotated ranking. The resultant factual density rankings of the two data sets in the first and second instance are given in Table 1a and Table 1b, and Table 2a and Table 2b respectively (see Appendix). Following a similar statistical approach as taken in [5], Table 3 provides with details on Spearman’s ρ rank correlation coefficient between the two data sets in both of the evaluation instances, as well as the respective values of the significance level. While [5] utilize the correlation outcome of the statistical analysis as to support its hypothesis on the adequacy of the factual density of the online web content relative to a human-annotated referent set, in context of our analysis the statistical analysis is used as to determine the feasibility of the automatic factual density calculation as informativeness metrics of the online news contents.

Table 3. Correlation tests on factual density ranking of online news articles using spearman’s rank correlation coefficient

Confidence Threshold	Spearman’s ρ	p-value	Significance Level
n/a	-0.013	0.471	52.9%
0.6	-0.235	0.094	90.6%

5. THE APPROPRIATENESS OF THE FACTUAL DENSITY AS AN INFORMATIVENESS MEASUREMENT FOR ONLINE NEWS

As Table 3 suggests, the factual density is far from an appropriate informativeness measurement for online reports of worldwide public interest. Given the enormous attention devoted to the investigation progress of the missing Malaysia Airlines Flight 370, an intuitive conjecture is to expect that most of the articles will yield with a factual set that can easily be utilized as a relevant and highly informative source on the overall incident. However, the general comparison suggests that this is not the case, questioning the impartiality and the reliability of the online news as an “influential source for making informed decisions” [2], [3], as well as the actual appropriateness of the ReVerb Open IE tool used in the factual count calculation.

Nevertheless, reconsidering the general factual density derivation effort in terms of increased confidence score on the factual counting brings presumably sufficient corroborating evidence on the adequacy of the factual density as an informativeness measurement of online news articles. The considerable improvement in the correlation coefficient – although still considered weak – brings the value of the significance level acceptable to a certain degree relative to the task of utilizing the online news as a referent source of information. Again, the values obtained in the analytical instance considering the confidence threshold of 0.6 by no means eliminate the general conclusion outlined above, suggesting that despite the fact that the factual density can be appropriately employed for the purpose of quality assessment under certain conditions, it still has to be taken with a degree of caution considering the inherit threat of journalistic bias.

In this context, it is also interesting to consider the rankings relative to utilization of the news articles in building a chronological overall report of the incident. Relying only on the assumed discriminative power of the factual density over the quality of a given online news recourse, the comparison of the rankings outlined in Table 2a and Table 2b suggests an unbalanced potential set of chronological references, having the Fox News articles dominating the incident briefing at the period of the very beginning and the critical end of flight data recorder battery lifetime, leaving the Thompson-Reuters articles to dominate the other portion of the chronological

reporting with a slightly better level of informativeness. Aware of the fact that behind this particular outcome various different causal relationships can be identified, it was brought to the discussion more from the information assurance and security perspective on the utilization of the informativeness metrics. As such, it is worth mentioning that the informativeness – measured using the factual density - can easily be manipulated and exploited as to produce a different, but equally plausible, set of chronological summaries of the incident by compiling daily reports to contain a high number of potentially extractable facts. Having *a priori* favourable fact count, one can easily modify the article to include facts that can lead to inferences in the final chronological report pre-programmed to correspond to its malicious intentions. Furthering the argument towards the actual appropriateness of the ReVerb tool for an informativeness assessment, the overall findings also suggest that despite the claim that ReVerb fails to utilize the semantic dimension in order to improve the robustness of the factual extraction, since no significant resilience level of the factual density can be recognized according to the overall analytical outcome.

6. CONCLUSION

The work presented in this paper reveals interesting aspects on the appropriateness of the applicability of the factual density measurement in the attempt for an automatic derivation of the online news articles quality in terms of informativeness. Considering the nature of the online journalism together with the limitations of the factual determination relative to the involvement of a strong semantic component yield to a general conclusion that the informativeness assessment is valid under strict conditions of increased confidence level of the automatic factual extraction. Since the analytical effort focused on evaluation only on the automatically-derived factual density ranking, a plausible future direction of the appropriateness assessment is to compare these results against human-annotated ranking as to determine the overall usability of the online news content, at least in case of massive event reporting. Notwithstanding the useful informativeness included within the analysed reports, a year after the incident the Malaysian Airlines Flight 370 is still missing without any optimistic lead on critical evidence – including the flight data recorder - that can possibly shed light on the overall investigation.

APPENDIX

Table 1a. Factual density ranking of the Fox News articles without the confidence score

Date	Fact Count	Word Count	Factual Density $f_d(t)$
3/8/2014	69	1345	0.051301115
3/9/2014	10	203	0.049261084
3/10/2014	70	992	0.070564516
3/11/2014	44	778	0.05655527
3/12/2014	67	1,015	0.066009852
3/13/2014	66	1,039	0.063522618
3/14/2014	44	678	0.064896755
3/15/2014	6	130	0.046153846
3/16/2014	73	1,104	0.066123188
3/17/2014	85	1,448	0.058701657
3/18/2014	15	146	0.102739726
3/19/2014	27	897	0.030100334
3/20/2014	87	1,177	0.073916737
3/21/2014	45	787	0.057179161
3/22/2014	68	1,090	0.062385321
3/23/2014	82	1,263	0.064924782

3/24/2014	72	1,150	0.062608696
3/25/2014	67	1139	0.058823529
3/26/2014	78	1315	0.059315589
3/27/2014	69	1088	0.063419118
3/28/2014	70	1170	0.05982906
3/29/2014	7	113	0.061946903
3/30/2014	67	1434	0.046722455
3/31/2014	75	989	0.075834176
4/1/2014	50	828	0.060386473
4/2/2014	73	1170	0.062393162
4/3/2014	60	1067	0.056232427
4/4/2014	20	497	0.040241449
4/5/2014	45	497	0.09054326
4/6/2014	13	283	0.045936396
4/7/2014	73	1124	0.064946619
4/8/2014	42	823	0.051032807
4/9/2014	35	619	0.056542811

Table 1b. Factual density ranking of the Thomson Reuters articles without the confidence score

Date	Fact Count	Word Count	Factual Density $f_d(t)$
3/8/2014	64	979	0.065372829
3/9/2014	30	518	0.057915058
3/10/2014	95	1,196	0.079431438
3/11/2014	70	958	0.073068894
3/12/2014	10	724	0.013812155
3/13/2014	66	941	0.070138151
3/14/2014	43	666	0.064564565
3/15/2014	85	1,253	0.067837191
3/16/2014	90	1,346	0.066864785
3/17/2014	25	517	0.048355899
3/18/2014	70	996	0.070281124
3/19/2014	72	1033	0.069699903
3/20/2014	63	1,020	0.061764706
3/21/2014	62	991	0.062563068
3/22/2014	72	1,122	0.064171123
3/23/2014	50	803	0.062266501
3/24/2014	63	903	0.069767442
3/25/2014	37	655	0.05648855
3/26/2014	71	1091	0.06507791
3/27/2014	18	340	0.052941176
3/28/2014	55	975	0.056410256
3/29/2014	36	680	0.052941176
3/30/2014	46	726	0.063360882
3/31/2014	22	418	0.052631579
4/1/2014	39	812	0.048029557
4/2/2014	45	696	0.064655172
4/3/2014	24	409	0.058679707
4/4/2014	49	712	0.068820225
4/5/2014	52	778	0.066838046
4/6/2014	50	732	0.068306011
4/7/2014	45	669	0.067264574
4/8/2014	50	935	0.053475936
4/9/2014	55	1007	0.054617676

Table 2a. Factual density ranking of the Fox News articles with a 0.6 confidence score

Date	Fact Count	Word Count	Factual Density $f_d(t)$
3/8/2014	35	816	0.042892157
3/9/2014	3	79	0.037974684
3/10/2014	25	802	0.03117207
3/11/2014	15	313	0.047923323
3/12/2014	20	578	0.034602076
3/13/2014	16	473	0.033826638
3/14/2014	13	449	0.028953229
3/15/2014	2	59	0.033898305
3/16/2014	25	750	0.033333333
3/17/2014	30	927	0.03236246
3/18/2014	0	0	0
3/19/2014	7	190	0.036842105
3/20/2014	32	799	0.040050063
3/21/2014	45	787	0.057179161
3/22/2014	21	576	0.036458333
3/23/2014	31	980	0.031632653
3/24/2014	22	671	0.032786885
3/25/2014	19	710	0.026760563
3/26/2014	31	999	0.031031031
3/27/2014	22	686	0.032069971
3/28/2014	24	563	0.042628774
3/29/2014	1	12	0.083333333
3/30/2014	24	850	0.028235294
3/31/2014	24	374	0.064171123
4/1/2014	16	487	0.032854209
4/2/2014	27	825	0.032727273
4/3/2014	25	715	0.034965035
4/4/2014	14	379	0.036939314
4/5/2014	14	391	0.035805627
4/6/2014	5	145	0.034482759
4/7/2014	15	397	0.037783375
4/8/2014	13	341	0.038123167
4/9/2014	12	207	0.057971014

Table 2b. Factual density ranking of the Thomson Reuters articles with a 0.6 confidence score

Date	Fact Count	Word Count	Factual Density $f_d(t)$
3/8/2014	33	829	0.039806996
3/9/2014	9	225	0.04
3/10/2014	68	1,238	0.054927302
3/11/2014	29	936	0.030982906
3/12/2014	4	45	0.088888889
3/13/2014	32	838	0.038186158
3/14/2014	9	291	0.030927835
3/15/2014	29	786	0.036895674
3/16/2014	32	789	0.040557668
3/17/2014	10	240	0.041666667
3/18/2014	27	640	0.0421875
3/19/2014	22	754	0.029177719
3/20/2014	24	632	0.037974684

3/21/2014	21	593	0.035413153
3/22/2014	37	776	0.047680412
3/23/2014	25	523	0.047801147
3/24/2014	24	609	0.039408867
3/25/2014	16	335	0.047761194
3/26/2014	31	713	0.043478261
3/27/2014	10	216	0.046296296
3/28/2014	19	511	0.037181996
3/29/2014	23	507	0.045364892
3/30/2014	22	492	0.044715447
3/31/2014	11	226	0.048672566
4/1/2014	19	501	0.037924152
4/2/2014	16	368	0.043478261
4/3/2014	10	228	0.043859649
4/4/2014	19	442	0.042986425
4/5/2014	18	546	0.032967033
4/6/2014	18	518	0.034749035
4/7/2014	17	491	0.034623218
4/8/2014	17	525	0.032380952
4/9/2014	18	410	0.043902439

REFERENCES

- [1] Nielsen, "111 Million Americans go on the Web for Daily News," Newswire, 2013. [Online]. Available: <http://www.nielsen.com/us/en/newswire/2013/todays-top-headline-111-million-americans-go-on-the-web-for.html>.
- [2] Fleishman Hillard, "Digital Influence Index 2012," London, UK, 2012.
- [3] Reuters Institute - Oxford University, "Digital News Report 2013," Oxford, UK, Jun. 2013.
- [4] E. Lex, M. Voelske, M. Errecalde, E. Ferretti, L. Cagnina, C. Horn, and M. Granitzer, "Measuring the Quality of Web Content using Factual Information," in Proceedings of the 2nd Joint WICOW/AIRWeb Workshop on Web Quality, 2012, no. iii, pp. 7–10.
- [5] C. Horn, A. Zhila, A. Gelbukh, R. Kern, and E. Lex, "Using Factual Density to Measure Informativeness of Web Documents," in In Proceedings of the 19th Nordic Conference on Computational Linguistics (NODALIDA), 2013.
- [6] E. Lex, A. Juffinger, and M. Granitzer, "Objectivity classification in online media," in Proceedings of the 21st ACM conference on Hypertext and hypermedia - HT '10, 2010, pp. 293–294.
- [7] N. Weber, K. Schoefegger, T. Ley, S. Lindstaedt, A. Brown, and S. Barnes, "Knowledge Maturing in the Semantic MediaWiki : A Design Study in Career Guidance," Lect. Notes Comput. Sci., vol. 5794, pp. 700–705, 2009.
- [8] J. E. Blumenstock, "Size Matters : Word Count as a Measure of Quality on Wikipedia," in Proceedings of the 17th international conference on World Wide Web, 2008, pp. 1095–1096.
- [9] S. Neuman, "Search For Flight MH370 Reportedly Largest In History," NPR, 2014. [Online]. Available: <http://www.npr.org/blogs/thetwo-way/2014/03/17/290890377/search-for-flight-mh370-reportedly-largest-in-history>.
- [10] A. Fader, S. Soderland, and O. Etzioni, "Identifying Relations for Open Information Extraction," in EMNLP '11 Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2011, pp. 1535–1545.
- [11] D. S. Etzioni, O., Banko, M., Soderland, S., & Weld, "Open Information Extraction from the Web," Commun. ACM, vol. 51, no. 12, pp. 68–74, 2008.
- [12] P. Gamallo, M. Garcia, and S. Fern, "Dependency-Based Open Information Extraction," in Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, 2012, pp. 10–18.
- [13] F. Wu and D. S. Weld, "Open Information Extraction using Wikipedia," in Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 2010, pp. 118–127.

- [14] D. Irvine, “Cable News Ratings: Fox Still Number One—MSNBC, CNN Shedding Viewers,” Accuracy in Media, 2014. [Online]. Available: <http://www.aim.org/don-irvine-blog/cable-news-ratings-fox-still-number-one-msnbc-cnn-shedding-viewers/>.
- [15] R. Sambrook, “Delivering trust: impartiality and objectivity in the digital age,” 2012.
- [16] Investigations Office of Aviation Safety, “Flight Data Recorder Handbook for Aviation Accident,” 2002.
- [17] F. Wu and D. S. Weld, “Autonomously semantifying wikipedia,” in Proceedings of the sixteenth ACM conference on Conference on information and knowledge management - CIKM '07, 2007, pp. 41–50.

AUTHORS

Filipo Sharevski (born 1985) is a Ph.D. Candidate at the Purdue University in the domain of Cybersecurity. In 2008 and 2011 he earned the B.Sc. and M.Sc. degrees in Telecommunications from the Faculty of Electrical Engineering and Information Technologies at University “Ss. Cyril and Methodius” in Skopje, Macedonia. After graduation, he was affiliated with the Institute of Telecommunications at the Faculty of Electrical Engineering and Information Technologies, Vip Macedonia Mobile Network Operator as networking engineer, and with Purdue University as a teaching and research assistant in the domain of Cybersecurity. His research interests are primarily focused on cyber forensics, network security, advanced mobile communication technologies and cyber warfare, as well as the natural language information assurance and security.

