Retrieval of Visually Shared News

Dmitrijs Milajevs

Guest Researcher at NIST, Maryland, USA
dimazest@gmail.com

Abstract. Some state funded media outlets are known to disseminate disinformation. The scientific community has shown a lot of interest in the network of bots that amplify this content. This work studies how the content of media outlets is spread among individual users on social media. We focus on tweets that include news articles as visual attachments because they are very powerful in delivering a convincing message, but are hard to retrieve. We propose and evaluate a model that is able to identify screenshots of news articles with a precision of 0.23 and a recall of 0.80. As an demonstration of the model, we analyzed the tweets of Twitter users from Latvia and identified patterns of news distribution.

Keywords: Social media · News · Opinion Manipulation.

1 Introduction

With the growing adoption of social media, the internet, originally an academic project, has become an integral part of everyday life experience. This shift attracted adversarial practices such as user tracking for targeted advertisement and opinion manipulation for political and personal goals [8]. Concerns that manipulation of public opinion by foreign countries affects the quality of public policy and democracy are widely promulgated in the press and media manipulation is studied by academics [15,11]. While the main forces behind opinion manipulation are clear—disinformation and polarization—it has yet to be established to what degree they reach their ultimate goals [19].

As the internet is a global network, opinion manipulation is a global phenomenon [17]. The Brexit referendum in the UK, the presidential election in the US and elections in Europe are stark and worrying examples where networks of automated bots often performed the task of distributing content of support and discredit [20]. A study of the computational propaganda phenomena in the conflict in Ukraine showed that the Twitter accounts of Russian mass media actively disseminated disinformation [21,5]. Techniques and strategies are being developed to combat Russian propaganda in Eastern and Central Europe [10,9]. In a twist of circumstance, while the Internet was initially believed to be a vehicle that would spread Western ideals around the world, it seems currently more likely to be viewed as a threat to these same ideals [13].

To understand how manipulation is done, [4] looked at narratives about Nordic-Baltic countries promoted by Russian media by analyzing three Russian websites and identifying dominating narratives. To assess the influence of those narratives, they performed a public opinion survey. [18] examined the frames—the packaging and the presentation of ideas that determine how information is processed and acted upon—of the narratives spread in social media by Russian propaganda to build distrust in the US justice system.

To understand how public opinion is manipulated to affect public policy, in addition to knowing the narratives that are spread by propagandists, the way the narratives are framed, and the amplifying mechanisms, it is important to understand how the general public is exposed to artifacts of opinion manipulation. On Twitter, news are shared not only as links and text-based citations, but also as screenshots. Opinion manipulators can potentially abuse this functionality by reframing the content of an article according to their agenda.

We focus on the performance of a Russian state funded news agency Sputnik that is involved in disinformation. We extend our examination to all media outlets that draw attention of Twitter users from Latvia to detect other manipualtors. Media outlets from other countries serve as points of comparison.

Our data collection procedure is user-centric and neither media-centric nor bot-centric. Instead of analyzing content produced by mass media or bots in isolation, we attempt to base our analysis on the data that is representative of the cyber environment of a country. As a result, we can contrast the content produced by the media with its reflection in public discourse.

The notion of the Latvian cyber space is central to this study. Section 2 describes the method that was used to identify Twitter accounts from Latvia. It is based on culturally distinctive keywords. In this case, names of a nationwide celebration are used. Over a thousand individual, media and governmental accounts from Latvia were successfully identified. The analysis of conventional Twitter statistics—number of tweets, retweets, mentions and URL shares—shows that Sputnik actively publishes content in Latvia, but receives very little engagement. However, Sputnik's content appears in a more obscure way as visual attachments. Such engagement is difficult to recognize and measure.

Section 3 proposes and evaluates a model for recognition of mass media content shared visually. A vast amount of shares of mass media content that is difficult to measure was identified. Section 4 analyzes it to shed light on how mass media content in visual form is shared. Three different patterns are discussed in Section 4: 1. discussion of materials of Latvian media is mostly constructive; 2. discussions that originate from materials by Russian media or about Russia are provocative; 3. Western media is a mix between the two. In addition to the sentimental differences, Latvian Twitter users reach media through various sources. Latvian media is shared by Latvian users. Russian media shares originate mostly from either Russian or Ukrainian users. Western media reaches Latvian users through local and foreign shares.

2 Culturally distinctive keywords

To build a user-centric collection, a two stage procedure was employed: 1. identify Twitter accounts that belong to the Latvian segment of the internet and 2. collect tweets of those accounts for further analysis.

First stage The criteria used to identify a Twitter account as belonging to the Latvian Twittersphere was whether they tweeted about the Midsummer celebration. Midsummer is celebrated in Latvia during the 23rd and 24th of June. The celebration days are referred to as the Līgo day and the Jāņi day respectively. In Russian, the holiday is commonly referred to as Лиго. We queried the Twitter streaming API with these three words during the period of June 10–28 2018 to obtain our initial data.[1]

During the first stage, 3 141 tweets were collected. After a manual inspection, only 24 tweets were found not directly related to the celebration. The unrelated tweets were either produced by the trending hashtag reporting bots[2] or matched the query because of a character overlap: the surname Стерлигов and the word полигон (a military testing ground) both contain the character string лиго. Accidentally, the tweets matched by the erroneous character overlap were about events in Ukraine.

Second stage The tweets collected during the first stage were produced by 1 863 accounts. During the period of July 29 to November 30 2018, we queried the Twitter Streaming API again using the `follow` parameter for the tweets that were produced by these accounts or that mentioned them.[3] The second stage resulted in the final collection of 3 932 984 tweets.

2.1 Annotation

To asses the effectiveness of the account selection procedure, all 1 863 accounts were manually annotated by two aspects: the perceived location of an account and its type.

To highlight that **locations** refer to the cyber world, they were abbreviated using top level domains, for example, `.lv` for Latvia, `.ua` for Ukraine, `N/A` was used when it was not clear what location an account belonged to.

The other aspect is the **category** of an account (individual, commercial, collective, governmental or media). No effort was made to distinguish real accounts from bots that pretended to be a real person. We defined the difference between

[1] An English transliteration, for example *Ligo*, was not used, as the English term is ambiguous: it also stands for The Laser Interferometer Gravitational-Wave Observatory (LIGO) among other meanings. We also have noticed that it is common to use Latvian spelling in English tweets.

[2] Most of the time, the bots were listing national teams playing at the FIFA World Cup, together with Jāņi or Līgo, as they were trending in Latvia during that time.

[3] Due to technical reasons, the tweets from August 20 to September 5 were not collected and are not present in the final collection.

individual and collective accounts as an individual account would use the pronoun *I*, while a collective account would use *we*. Non-profit organizations, music bands and schools are examples of collective accounts.

2.2 Retrieval analysis

As tweets were labeled by the language they were written in, we employed that information as well. Twitter uses a combination of the ISO 639-1 alpha-2 and the ISO 639-3 alpha-3 codes. For example, `lv` stands for Latvian and `uk` stands for Ukrainian. The combination of the tweet level language annotation and the user level location annotation gave insight into the collection.

Location As expected, the majority of accounts were from Latvia (1 586, 85%). The locations of 135 accounts (7%) were unidentified. For the countries that were represented by more than 10 accounts, 37 (2%) accounts were from Russia, where Midsummer is celebrated by Latvians who have been deported to Siberia and their relatives. Unexpectedly, 35 (2%) accounts were from Ukraine. Such a high number cannot be explained neither by erroneously matched tweets in the first stage (those users produced only 16 467 tweets, which is 4% of all tweets that were created by the accounts from Ukraine), nor by cultural ties with Latvia, (there were only 6 and 3 accounts from the neighboring Baltic states, Lithuania and Estonia, respectively). 17 accounts (1%) came from the US, which can be explained by the presence of a Latvian diaspora in the country.

Language There were more tweets in Russian (27%) than in English (12%), which is not in line with a previous study. [12] built a tweet collection following accounts of Latvian media, government and public figures. In that collection, 74% of tweets were in Latvian, 10% in Russian, 10% in English and 6% were in other languages. The location of the accounts revealed that most of the tweets in Russian were not created by accounts from Latvia. For a large portion of these tweets, the account location could not be identified. Accounts from Russia, Ukraine and Latvia almost equally contributed to the tweets in Russian.[4]

We were unable to identify a location of a large proportion of accounts tweeting in Russian because they did not explicitly specify their locations or the locations of the accounts were unclear from the content of their tweets. For instance, some accounts simply posted critical comments on Russian politics and could either be from Russia, Ukraine or other country. The language distribution of tweets of users from Latvia is similar to the distribution reported in [12]: most of tweets are in Latvian (498 654, 75%), 53 582 tweets (8%) are in Russian and 44 207 tweets (7%) are in English.

It is still unclear why initially there were more Russian tweets retrieved than English. Majority of tweets in Russian came from 207 (11%) users from Russia,

[4] Because by location we mean *perceived* location, it does not mean that a Ukrainian account is actually a legitimate account from Ukraine. It could be a bot operated by a foreign actor. Because of this, statements about not Latvian accounts are not designed to be representative of the whole population of the corresponding group.

Ukraine or an undefined location. These 11% of users produced 31% of all tweets and 75% of tweets in Russian. We discuss in the next section how these active users got interested in a local Latvian celebration.

2.3 Anomalies due to Midsummer coverage by Sputnik

Among the tweets collected in the first stage (the tweets that mentioned the Midsummer celebration) there were two tweets that referred to the same content published by a Russian news agency Sputnik. Both of these tweets were actively retweeted: we captured 28 and 41 retweets.

Sputnik's tweets As the Midsummer celebration approached, Sputnik published various articles on its websites accompanied by corresponding announcements by several Twitter accounts.

Two tweets were posted on June 20 and came from the Latvian branch. They were two translations of the same material about a survey on Midsummer celebration traditions. The study was performed by a grocery store. Both articles linked to the same post on the website of the store and could be promotional material. On June 21, an account of Sputnik Lithuania released an excerpt of an audio recording of an interview about the Midsummer celebration in Latvia. The accompanying text focused on an attempt by the Latvian political elites to tie the celebration with politics. On June 22, Sputnik Latvia posted an article on the history of Midsummer celebration. A brief search of similar articles revealed two almost identical copies. One of the copies was announced on Twitter. On June 23, Sputnik Latvia published an article about an art exhibition dedicated to the Midsummer. On June 24, Sputnik Latvia posted pictures from the Midsummer celebrations in Riga, noting that it was visited by a former president of Latvia Vaira Vīķe-Freiberga.

Tweets that referred to Sputnik's content Sputnik-published content was also referenced by three individual accounts. All of them refered to the same material published on Sputnik Latvia's website. The material was essentially the same audio excerpt of the interview published by Sputnik Lithuania. However, the accompanying text focused on the connection, expressed in the interview, between Neo-Nazism and the Midsummer celebrations in Latvia and Ukraine. The more aggressive framing of these tweets attracted much more attention on Twitter and tweets referencing this material received likes and got retweeted (refer to Figure 1). The engagement grew enough that it attracted the attention of the Museum of the Occupation of Latvia, which subsequently debunked the interview.

The museum was not alone in resisting the spread of the narrative that Nazism is spreading in the Baltics and Ukraine (a similar narrative was identified in [4]). The tweets that referenced to the interview published by Sputnik pointed out its propagandistic nature by articulating that Russian propaganda links virtually anything to Nazism.

Tweet text: Squawk! And The St. John day, the Kupala Night, Līgo (Jāņi) in Latvia, becomes "fashist." A wonderful discovery! In this manner, soon they will get to Maslenitsa [the Butter Week, the Crepe week]. Because, a crepe is also associated with the Sun, the same as the Swastika does. Hitler's party has devoted a great deal of attention to such traditions. [The Face With Tears of Joy emoji].
Article title: Gasparyan: Līgo is interesting not because of the grass market, but because of neo-Nazis.
Article text: A historian, a political expert Armen Gasparyan believes, that Midsummer, which is being celebrated in Latvia already starting from today, is interesting neither because of a grass market nor festive celebrations.
"For a very long time, Midsummer became associated in the Baltics, as in Ukraine, with neo-Nazis. They did not come up with this in Riga, but it is taken from the Post-Germanic tradition, from the very beginning Hitler's party put a lot of attention to such traditions", — Gasparyan told in an interview to Sputnik Latvia.
Given that, according to him, the situation with the neo-Nazis will not get out of control.
Account description: The Lithuanian junta/applied Russophobia and breach of the sanctity of the Vatniks [literally: people who wears quilt jackets; a reference to people who are loyal to Putin.]/He only earns his freedom and existence, Who daily conquers them anew/karatel.foss.org.ua

Fig. 1. An exaple of a tweet posted in Russian that contains a news article as a screenshot.

The missing link to Ukrainian users Two authors of tweets that discussed Sputnik articles were identified as individual accounts from Latvia as they mostly tweeted in Latvian. The author of the tweet in Figure 1 was identified as an individual account from Lithuania who mostly tweeted in Russian about Russia, Ukraine, and the Baltics. As Sputnik content attracts international attention, we examined the frequency and location of accounts that engaged with Sputnik's content about the Midsummer tweet.

For Ukraine, 21 811 tweets (37%) came from 22 (63%) users who retweeted at least one of the tweets that discussed Sputnik's articles.

The account from Lithuania had a wide audience in Ukraine, thus a lot of users from Ukraine who retweeted his Midsummer tweet were identified as potential users from Latvia. Presumably, to attract attention, the user of this account described itself, among other things, as practicing "applied Russophobia," which explains the general anti-Russian sentiment of Russian tweets in the collection. Furthermore, its wide usage of irony suggests that it might be a trolling account.

Sputnik's influence Sputnik's own tweets might have not attracted much attention directly because their underlying material was of low quality. These included promotional materials and reposts. So far, no support of Sputnik's content was found on Twitter in the form of retweets leading to a positive discussion. Even though same content was discussed elsewhere: the article on the Midsummer celebration history that was posted on a blog-publishing service imhoclub.lv received 241 comments from 25 users while the same text published on Sputnik's website did not spark any discussion at all.

Sputnik received strong criticism of being an outlet of Russian propaganda and was sarcastically mocked by some Twitter users from Latvia. However, not all of Sputnik's content was labeled as propagandistic. For example, the aforementioned historical review was not called out. Neither was the interview posted by Sputnik Lithuania, but the same interview with a more aggressive description on Sputnik Latvia was.

Content published by Sputnik in Russian was labeled as propaganda not only in Russian, but also in Latvian.

Material deemed to be propaganda was called out in different ways. Providing a link to the material was one method used. It is the clearest method, as both the source is known and the context is given by the Twitter user interface. Taking a screenshot of the source article was another way as shown in Figure 1. It is less clear, as only the context is given, while the source needs to be inferred indirectly either by a web search or design elements.

Even though Sputnik did not get a lot of attention, it operated in an agile manner by publishing a vast amount of material varying presentation and framing [18,4]. Counteraction creates a feedback loop revealing the techniques that are effective in gaining attention of users from Latvia [11]. Despite low overall performance this knowledge can be used by manipulators to effectively reach Latvian demographics when needed.

3 Identification of visually shared mass media content

We have found so far that individual Twitter accounts from Latvia do not actively engage with Sputnik's content by retweets, mentions or URL shares. However, does Sputnik's content appear in more obscure ways? One of the methods employed on Twitter to share mass media content such as a news article, is to share it as a visual attachment, for example as a screenshot of its relevant part. The image gives the context, while the text of a tweet provides the message as it is shown in Figure 1.

There are many ways for mass media content to be shared visually. A photo of an article from a newspaper can be taken with a smartphone. Alternatively, a screenshot of the same article can be taken from the newspaper's website. The composition of a visual attachment varies as well: the attachment can include only the text of the body of the article or only a headline with a photo. Design elements such as a website logo and other distinctive elements help to identify the source, but they are not always present.

All these factors already make it difficult to come up with a clear set of identification rules. The situation becomes even more complicated when there exist visual attachments that share similarities with images of news articles that are not actually such images. For example, images that contain quotes or screenshots of books appear similar to news articles. Even though it is difficult to formally characterize visually shared mass media content, it is nevertheless easy for humans to identify this content. Such task is a good fit for an approach based on machine learning which is able to identify necessary patterns in the data to perform classification without feature engineering. A method based on machine learning requires labeled data: positive and negative examples of the screeshots of news articles. This section provides details on the dataset and the model we used for identifying visually shared mass media content. First, the dataset is described. Later, the model is introduced and evaluated.

3.1 The image dataset

As we were interested in the behavior of Twitter users from Latvia, we used the tweets and the retweets produced by individual users from Latvia discussed in Section 2. As our **test data**, we randomly sampled 19 877 tweets (16% of 124 864 tweets that were produced by individual users from Latvia and contained an image) and found that 644 (3%) of them contained news article screenshots.

In an attempt to increase the proportion of positive samples, we sampled[5] 11 391 tweets from the users who posted at least one tweet with a picture of mass media content in the test data. Among the subset of these tweets, 633 (6%) tweets were positive. These tweets were used as **validation data** to select the best candidate model.

Finally, for **training data**, we labeled all the tweets produced by users who posted at least 6 tweets with pictures of mass media content in validation data. This resulted in 18 910 tweets, 1 176 (6%) of which were labeled positively.

At this point, the duplicates across and within the datasets were not removed. Retweets are one source of duplicates, but we also noticed that sometimes the same picture was purposely reused in different tweets, especially when a tweet contained several images.

3.2 The model

We designed a model with a stack of three convolutional layers each followed by a max-pooling layer. We used two inter-connected layers with a dropout layer in between on top. To support classification of images of different sizes, a global max-pooling layer was placed between the convolutional layers and the inter-connected layers. The output of the first inter-connected layer was used to get image embeddings. The second layer was a single layer with sigmoid activation for a binary classification.The model was implemented using `Keras` 2.2.4 [6] with `TensorFlow` 1.12.0 [1] as the backend, `CUDA` 10.0 and `cuDNN` 7.5.

This model architecture introduced some implementation choices which we expressed in the form of hyper-parameters. We parametrized the number of filters and filter size of convolutional layers: the first two layers shared the hyper-parameter values, the third layer had its own hyper-parameter values. The pool

[5] These steps were not performed strictly one after another, instead there were several iterations.

[6] Refer to Keras documentation for the optimizer descriptions.

Table 1. Hyper-parameters. Values in bold are the values of the manually chosen model.

Hyper-parameter	Values
Convolution layers 1 and 2 filter number	8, 16, **32**, 64
Convolution layers 1 and 2 kernel size	**5**, 7, 9, 11
Max-pooling layers 1 and 2 pool size	2, **4**
Convolution layer 3 filter number	8, **16**, 32, 64
Convolution layer 3 kernel size	5, 7, 9, **11**
Max-pooling layer 3 pool size	**2**, 4
Embedding size	**16**, 32, 64, 128
Dropout	0.1, 0.2, 0.3, 0.4, **0.5**, 0.6, 0.7, 0.8, 0.9
Epochs	**50**, 100, 150, 200, 250, 300
Optimizer[6]	sgd, rmsprop, **adagrad**, adadelta, adam, adamax, nadam

size of max-pooling layers was similarly parametrized. The embedding size, or the size of the first inter-connected layer, was also parametrized. Finally, we experimented with several optimizers and numbers of epochs during training. Table 1 summarizes the hyper-parameters and lists their possible values. The rectified linear unit (ReLU) was used as an activation function.

3.3 Model training and hyper-parameter tuning

Candidate models were trained on the training dataset and evaluated on the evaluation dataset. The evaluation and optimization metric was the F_2 score [16]. The F_2 score was chosen to make recall more influential than precision, but to take both values into account. Binary cross-entropy was used as the loss function.

A Bayesian optimization method was employed to tune the hyper-parameters. The Tree of Parzen Estimators algorithm (TPE, [2,3]), implemented in `hyperopt` 0.1.1, was used together with random search and annealing algorithms. The algorithms were randomly chosen at each iteration, giving TPE a preference of 0.7, annealing 0.2 and random search 0.1. Such a mixture was used to avoid bias in hyper-parameter tuning.

The model selection process was run twice (an initial run to remove duplicates, and a deduplicated run) for 350 iterations in the first run and 1 000 iterations in the second. Models were trained in parallel on the NIST Enki HPC cluster[7] on 13 nodes, each equipped with 4 Nvidia Tesla V100 SXM2 GPUs.

Three early-stopping conditions during model training were used: training would stop if the accuracy dropped lower than 0.6 for three epochs in a row or if it dropped lower than 0.8 after 50 epochs, or if it dropped lower than 0.9 after 100 epochs. Learning rate was decreased when model accuracy stopped improving by using `ReduceLROnPlateau()`, a callback provided in `Keras`.

As the images were different sizes, a batch size of 1 was used. Because the 3 pooling layers in the model reduced the dimensionality by a factor of 2 or 4

[7] `https://www.nist.gov/programs-projects/computation-platform-aiml`

Table 2. Model performance on dataset splits.

Split	Precision	Recall	F_2
Train	0.40	0.99	0.76
Validation	0.30	0.82	0.61
Test	0.23	0.80	0.53

each, there was a lower limit on the valid image size. Images smaller than 256 pixels in height or width were padded in either dimension to make sure that the smallest dimension of an image was at least 256 pixels.

Because there were many more negative samples than positive, during training the model tended to learn to identify any image as a negative instance, as it was unlikely that a positive sample was randomly picked. To mitigate this issue, positive instances were oversampled and negative instances were undersampled. Sampling was done without replacement. For every training epoch, a random positive instance followed a random negative instance. An epoch ended when all positive instances had been used.

3.4 The final deduplicated dataset

First, the model selection procedure (refer to Sections 3.2 and 3.3 above) was run over the data with duplicate images. The best model according to the model selection procedure was later used to identify duplicates. Concretely, the output of the second last inner-connected layer was used as a vector embedding of a corresponding image. The vectors were clustered using the DBSCAN algorithm [7] as implemented in `scikit-learn` version 0.20.2 [14] with the default settings, but the minimal number of samples to form a cluster was set to 2.

Once the clusters were identified, one sample per cluster was selected giving a preference first to a test sample, then to a validation sample and, lastly, to a training sample.

The deduplicated collection consists of 25 637 test, 9 072 validation and 11 242 training images. Regarding positive samples, there are 844 positive samples in test data, 537 in validation and 770 in training.

3.5 Evaluation

The large number of trained models during the model selection step led to a high spread in the precision and recall values. F_2 was the hyper-parameter optimization metric, but small differences in F_2 led to significant difference in precision and recall. The best F_2 score on the validation set was 0.62 and yielded precision of 0.43 and recall of 0.70. However, there were 21 models with an F_2 score greater than 0.60. For these models, the smallest precision was 0.29, the smallest recall was 0.65, the highest precision was 0.47, and the highest recall was 0.82.

Given the importance of recall, we manually selected a model with the highest recall among the models that achieved an F_2 score greater than 0.60. The manually selected model got an F_2 score of 0.61, precision of 0.30, and recall of 0.82. The drop by 0.12 recall points compared to the best model was not justified in the gain of 0.13 precision points, given that the output of the model would be manually inspected afterwards. We suspect that the F_2 score may have been a poor choice for the hyper-parameter optimization, as the small difference in 0.01 F_2 points led to a significant differences in precision and recall. However, we leave the study of a better performance metric for the future work.

The selected model achieved almost perfect recall of 0.99 on the training split, with a precision of 0.40, and an F_2 score of 0.76. As expected, the performance numbers on the validation split were lower, as was discussed above. On the test split, the model achieved precision of 0.23, recall of 0.80 and an F_2 score of 0.53. In comparison with the validation split, precision dropped, while recall did not change much, which is a favorable behavior as the downstream application is recall oriented. Table 2 summarizes these results.

Model's typical mistakes are the memes that contain a photo with some text, quotes of famous people and photos of politicians.

4 Visual shares

The previous section identified tweets that shared mass media content visually. This section looks how visual shares of local and foreign media outlets reached Twitter users from Latvia. Even though in Section 2 tweets from users from several countries were obtained, in this section we analyze only tweets of users from Latvia, though this also includes retweets of foreign tweets.

Tweet origin annotation Tweets were identified by location. For tweets that were not retweets the location of the author was used. For a retweet, the location of the user who created the tweet, not the location of the user who retweeted it was used. This is the extension of the annotation method performed in Section 2.1. However, geographically ambiguous accounts—accounts with locations that could not be distinguished as being from Russia or Ukraine—became more troublesome. The accounts were difficult to annotate mostly because they did not represent a person or a group of people and posted political commentary. For such cases, accounts were labeled as individual accounts from an undefined location.

Media origin annotation Along to the location attribution described previously, we also attributed locations to the mass media organizations whose screenshotted articles were tweeted and shared. The following location groups were used: **local** for Latvian media, **Baltic** for the media from Estonia and Lithuania, **Russian** for the media from Russia, **Eastern European** for the media from Eastern Europe (Ukraine is the major country in this category), **Western** for Western countries. When the origin of a screenshot was difficult to identify, it was labeled as **other**.

Table 3. Media origin of tweets that are tweeted or retweeted by users from Latvia.

Tweet origin	Local	Baltic	Russian	Eastern European	Western	Other
Latvia	2 328	11	185	9	423	47
Russia	1	-	93	6	21	6
Ukraine	-	-	59	15	29	10
United Kingdom	-	-	-	-	64	-
United States	2	-	7	-	125	1
N/A	4	-	294	12	117	36
Other	1	1	37	4	93	3

Fine-grained classification (for example, on a media outlet level) was avoided because some screenshots did not have clear references to the media outlet. Source attribution was also made harder by multiple copies of the same content hosted on several web sites, similarly to the article discussed in Section 2.3 on the Midsummer celebration that was posted by Sputnik and imhoclub.lv.

Table 3 shows the number of tweets and retweets of users from Latvia that contained mass media visual content. As seen in this table, most of Latvian media content was spread by users from Latvia and very few tweets came from foreign users. Twitter users from Latvia not only shared, discussed and expressed opinions about the articles of local media they shared, but also gave comments about the media outlets themselves. For example, Latvian users frequently tried to point out biased coverage or technical issues.

The majority of Russian media was mostly delivered through geographically ambiguous accounts along with accounts from Russia and Ukraine. Some of the Russian and Ukrainian accounts belonged to professional journalists or media outlets themselves. However, the majority of tweets were sarcastic political commentary authored by anonymous and pseudonymous accounts. Contrary to foreign accounts that picked news about Russia to back up their provocative messages, users from Latvia tended to share articles where Latvia and Latvians are mentioned. Consequently, when Latvia was used in propagandistic materials, discussions that stemmed from such shares were negative towards Russia.

Western media distribution pattern was the most diverse. Users from Latvia posted screenshots of Western media by itself and also retweeted tweets from foreign users that include Western media. A lot of content, especially posted by accounts from the US and the UK, comes from professional journalists. Their aim was to spread the news articles rather than commentary. Still, there were large numbers of sarcastically-framed shares that covered events related to Russia, such as the World Cup, the Skripal poisoning, and the Trump-Putin meeting in Helsinki.

5 Conclusion

This work was concerned with how the content of media outlets was shared by Twitter users from Latvia. First, we introduced our method of collecting a representative group of the population of a country. Our method was based on culturally important and globally unique celebration names. After constructing a tweet corpus, we identified some tweets with media shares to build a model that identified visual shares of media outlets. On the test dataset, the model recall was 0.80 and the precision was 0.23. The model was used to identify more tweets with media shares. Finally, media shares were then analyzed to shed light on how mass media content reaches Latvian Twitter users. This study identified three main mass media outlet groups that have varying distribution patterns and roles in overall Latvian Twitter discourse: Latvian media, Russian media, and Western media.

We found that media from Latvia almost exclusively reaches Latvian Twitter users via themselves. Latvian media is shared to either discuss the subject matter of the shared article or to criticize the media outlet. Normally, the discussions are constructive and engaging. Russian media mostly reaches Latvian users via foreign proxies. Most of these proxies are accounts that give political commentary about events related to Russia in a sarcastic manner. Western media plays an important role in the Latvian media environment. Western media originates from Latvian and foreign Twitter users. A lot of these visual media shares originate from journalists who work in the UK and the US. Western media outlets that publish content in Russian (Deutche Welle, BBC and Radio Free Europe) compete equally with the media from Russia.

This work identified several directions to be explored in the future. Though neural networks can be used in Information Retrieval, this work shows that the F_2 score is not precise enough to capture the importance of recall and it is not indicative for reporting the performance on the downstream task. This work was concerned with identification of mass media content, but the model can be extended to perform particular source identification. For example, the model can be extended to retrieve the screenshots of BBC articles. This can help with building a more precise tool to monitor media influence.

Latvian Twitter users are well aware of and sceptical about Russian propaganda. They actively counteract it. However just pointing out and making fun of propaganda is not an efficient method for counteracting it because it creates an optimal environment for trolls and extremists to exploit. Furthermore, it does not stop the original propagandistic message from being distributed. Effective counteraction needs to avoid the feedback loop to manipulators and propagandists.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng,

X.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), https://www.tensorflow.org/
2. Bergstra, J., Yamins, D., Cox, D.: Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: Dasgupta, S., McAllester, D. (eds.) Proceedings of the 30th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 28, pp. 115–123. PMLR, Atlanta, Georgia, USA (17–19 Jun 2013), http://proceedings.mlr.press/v28/bergstra13.html
3. Bergstra, J.S., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: Shawe-Taylor, J., Zemel, R.S., Bartlett, P.L., Pereira, F., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 24, pp. 2546–2554. Curran Associates, Inc. (2011), http://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization.pdf
4. Bērziņa, I., Cepurītis, M., Kaljula, D., Juurvee, I.: Russia's Footprint in the Nordic-Baltic Information Environment. Tech. rep., NATO Strategic Communications Centre of Excellence (2018), https://www.stratcomcoe.org/download/file/fid/78090
5. Bērziņš, J., Jaeski, A., Laity, M., Maliukevicius, N., Navys, A., Osborne, G., Tatham, S.: Analysis of Russia's information campaign against Ukraine. NATO StratCom Report (2014), https://www.stratcomcoe.org/download/file/fid/1886
6. Chollet, F., et al.: Keras. https://keras.io (2015)
7. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. pp. 226–231. AAAI Press (1996), https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf
8. Ghosh, D., Scott, B.: Digital Deceit: The Technologies Behind Precision Propaganda on the Internet. Tech. rep., New America (2018), https://www.newamerica.org/public-interest-technology/policy-papers/digitaldeceit/
9. Helmus, T.C., Bodine-Baron, E., Radin, A., Magnuson, M., Mendelsohn, J., Marcellino, W., Bega, A., Winkelman, Z.: Russian Social Media Influence: Understanding Russian Propaganda in Eastern Europe. RAND Corporation (2018), http://www.rand.org/t/RR2237
10. Lucas, E., Pomeranzev, P.: Winning the Information War: Techniques and Counter-strategies to Russian Propaganda in Central and Eastern Europe (2016), http://infowar.cepa.org/Winning-the-Information-War
11. Marwick, A., Lewis, R.: Media Manipulation and Disinformation Online. Tech. rep., Data&Society Research Institute (2017), https://datasociety.net/output/media-manipulation-and-disinfo-online/
12. Milajevs, D.: Language use in a multilingual tweet corpus. Frontiers in Artificial Intelligence and Applications **307**, 88–95 (2018). https://doi.org/10.3233/978-1-61499-912-6-88, http://doi.org/10.3233/978-1-61499-912-6-88, presented at the Human Language Technologies—The Baltic Perspective conference
13. Morozov, E.: The Net Delusion: The Dark Side of Internet Freedom. Perseus Books, Cambridge, MA, USA (2011)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

15. Phillips, W.: The Oxygen of Amplification: Better Practices for Reporting on Extremists, Antagonists, and Manipulators Online. Tech. rep., Data&Society Research Institute (2018), https://datasociety.net/output/oxygen-of-amplification/
16. Sasaki, Y., et al.: The truth of the f-measure. Teach Tutor mater **1**(5), 1–5 (2007)
17. Singer, P., Brooking, E.: LikeWar: The Weaponization of Social Media. Houghton Mifflin Harcourt (2018), https://www.likewarbook.com/
18. Spaulding, S., Nair, D., Nelson, A.: Beyond the Ballot: How the Kremlin Works to Undermine the U.S. Justice System. Tech. rep., Center for Strategic and International Studies (CSIS) (may 2019), https://www.csis.org/analysis/beyond-ballot-how-kremlin-works-undermine-us-justice-system
19. Tucker, J., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., Nyhan, B.: Social media, political polarization, and political disinformation: A review of the scientific literature. SSRN (Mar 2018), https://eprints.lse.ac.uk/87402/1/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf
20. Woolley, S.C., Howard, P.N.: Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media. Oxford University Press (2019), https://www.oii.ox.ac.uk/research/books/computational-propaganda-political-parties-politicians-and-political-manipulation-on-social-media/
21. Zhdanova, M., Orlova, D.: Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media, chap. Ukraine: External Threats and Internal Challenges. In: [20] (2019), https://www.oii.ox.ac.uk/research/books/computational-propaganda-political-parties-politicians-and-political-manipulation-on-social-media/