RES MILITARIS Social Science Journal

Social Media and Misleading Information in a Democracy A Mechanism Design Approach

G.N.V Vibhav Reddy[1],Varundath chowdary[2],M Mounya[3],J Bhargavi[4], G Pavan sai[5]

[1,2,3,4,5]Department of Computer Science and Engineering

[1,2,3,4,5] Sree Dattha Institute of Engineering and Science, Sheriguda, Telangana.

ABSTRACT

In this paper, we present a resource allocation mechanism for the problem of incentivizing filtering among a finite number of strategic social media platforms. We consider the presence of a strategic government and private knowledge of how misinformation affects the users of the social media platforms. Our proposed mechanism incentivizes social media platforms to filter misleading information efficiently, and thus indirectly prevents the spread of fake news. In particular, we design an economically inspired mechanism that strongly implements all generalized Nash equilibria for efficient filtering of misleading information in the induced game. We show that our mechanism is individually rational, budget balanced, while it has at least one equilibrium. Finally, we show that for quasi-concave utilities and constraints, our mechanism admits a generalized Nash equilibrium and implements a Pareto efficient solution.

Index Terms: Resource Allocation, Incentive Mechanism, Social Media Platforms, Misinformation, Fake News, Nash Equilibrium, Game Theory, Individual Rationality, Budget Balance, Pareto Efficiency.

1.INTRODUCTION

For the last few years, political commentators have been indicating that we live in a *post-truth* era [1], wherein the deluge of information available on the internet has made it extremely difficult to identify facts. As a result, individuals have developed a tendency to form their opinions based on the *believability* of presented information rather than its truthfulness [2]. This phenomenon is exacerbated by the business practices of social media platforms, which often seek to maximize the *engagement* of their users at all costs. In fact, the algorithms developed by platforms for this purpose often promote conspiracy theories among their users [3].

The sensitivity of users of social media platforms to conspiratorial ideas makes them an ideal terrain to conduct political misinformation campaigns [4], [5]. Such campaigns are especially effective tools to disrupt democratic institutions, because the functioning of stable democracies relies on *com- mon knowledge* about the political actors and the processes they can use to gain public support [6]. The trust held by the citizens of a democracy on common knowledge includes: (i) trust that all political actors act in good faith when contesting for power, (ii) trust that elections lead to a free and fair transfer of power between the political actors, and (iii) trust that democratic institutions ensure that elected officials wield their power in the best interest of the citizens. In contrast, citizens of democracies often have a *contested knowledge* regarding who should hold power and how they should use it [6]. The introduction of *alternative facts* can reduce the trust on common knowledge about democracy, especially if they become accepted beliefs among the citizens. Such disruptions on the trust on common knowledge can be found in the 2016 U.S. elections [7] and Brexit Campaign in 2016 [8], where the spread of misinformation through social media platforms resulted in a large number of citizens mistrusting the results of voting. To tackle this growing phenomenon of misinformation, in this paper, we consider a finite group of social media platforms, whose users represent the citizens in a democracy, and a democratic government. Every post in the platforms is associated with a parameter that captures its informativeness , which can take values between two extremes: (i) completely factual and (ii) complete misinformation. In our framework, posts that exhibit misinformation can lead to a decrease in trust on common knowledge among the users [9]–[12]. In addition, social media platforms are considered to have the technologies to *filter*, or label, posts that intend to sacrifice trust on common knowledge. Thus, the government seeks to incentivize the social media platforms to use these technologies and filter any misinformation included in the posts.

Motivated by capitalistic values, we induce a *misinformation filtering game* to describe the interactions between the social media platforms and the government. In this game, each platform acts as strategic player seeking to maximize their advertisement revenue from the engagement of their users [7], [13]. User engagement is a metric that can be used to quantify the interaction of users with a platform, and subsequently, how much time they spend on the platform. Recent efforts reported in the literature on misinformation in social media platforms have indicated that increasing filtering of misinformation leads to decreasing of user engagement [14]. There are many possible reasons for this phenomenon. First, filtering reduces the total number of posts propagating across the social network. Second, the users whose opinions are filtered may perceive this action as dictatorial censorship [15], and as a result, they may chose to express their opinions in other platforms. Finally, misinformation tends to elicit stronger reactions, e.g., surprise, joy, sadness, as compared to factual posts [16], which may increase user engagement. Thus, each platform is reluctant to filter misinformation.

In our framework, we consider that the government is also a strategic player, whose utility increases as the trust of the users of social media platforms on common knowledge increases. Consequently, increasing filtering of misinformation by the social media platforms increases the utility of the government. Thus the government is willing to make an investment to incentivize the social media platforms to filter misinformation. In our approach, we use mechanism design to distribute this investment among the platforms optimally, and in return, implement an optimal level of filtering.

Mechanism design was developed for the implementation of system-wide optimal solutions to problems involving multiple rational players with conflicting interests, each with private information about preferences [17]. Note that this approach is different from traditional approaches to decentralized control with private information [18]–[21] because the players are not a part of the same time, but in fact, have private and competitive utilities. The fact that Mechanism design optimizes the behaviour of competing players has led to broad applications spanning different fields including economics, politics, wireless networks, social networks, internet advertising, spectrum and bandwidth trading, logistics, supply chain, management, grid computing, and resource allocation problems in decentralized systems [22]–[28].

The contribution of this paper is as follows. We present an indirect mechanism to incentivize social media platforms to filter misleading information. We show that our proposed mechanism is (i) feasible, (ii) budget balanced, (iii) individual rational, and (iv) strongly implementable at the equilibria of the induced game. We prove the existence of at least one generalized Nash equilibrium and show that our mechanism induces a Pareto efficient equilibrium The rest of the paper is organized as follows. In Section II, we provide the modeling framework and problem formulation. In Section III, we present our mechanism, and in Section IV, we prove the associated properties of the mechanism. In Section V, we interpret the mechanism and present a descriptive example. Finally, in Section VI we conclude and present some directions for future research.

2.EXISTING SYSTEM

social media in particular, has generated extraordinary concern, in large part because of its potential effects on public opinion, political polarization, and ultimately democratic decision making. Recently, however, a handful of papers have argued that both the prevalence and consumption of "fake news" per se is extremely low compared with other types of news and news-relevant content. Although neither prevalence nor consumption is a direct measure of influence, this work suggests that proper understanding of misinformation and its effects requires a much broader view of the problem, encompassing biased and misleading—but not necessarily factually incorrect—information that is routinely produced or amplified by mainstream news organizations. In this paper, we propose an ambitious collective research agenda to measure the origins, nature, and prevalence of misinformation, broadly construed, as well as its impact on democracy. We also sketch out some illustrative examples of completed, ongoing, or planned research projects that contribute to this agenda.

Disadvantages

1) The system doesn't have facility to train and test on large number of numbers.
2) The system doesn't facility for analyzing the Nash-implementation.

3.PROPOSED SYSTEM

To tackle this growing phenomenon of misinformation, in this paper, we consider a finite group of social media platforms, whose users represent the citizens in a democracy, and a democratic government. Every post in the platforms is associated with a parameter that captures its informativeness, which can take values between two extremes: (i) completely factual and (ii) complete misinformation. In our framework, posts that exhibit misinformation can lead to a decrease in trust on common knowledge among the users [9]–[12]. In addition, social media platforms are considered to have the technologies to *filter*, or label, posts that intend to sacrifice trust on common knowledge. Thus, the government seeks to incentivize the social media platforms to use these technologies and filter any misinformation included in the posts.

In our framework, we consider that the government is also a strategic player, whose utility increases as the trust of the users of social media platforms on common knowledge increases. Consequently, increasing filtering of misinformation by the social media platforms increases the utility of the government. Thus the government is willing to make an investment to incentivize the social media platforms to filter misinformation.

In our approach, we use mechanism design to distribute this investment among the platforms optimally, and in return, implement an optimal level of filtering. Mechanism design was developed for the implementation of system-wide optimal solutions to problems involving multiple rational players with conflicting interests, each with private information about preferences [17]. Note that this approach is different from traditional approaches to decentralized control with private information [18]–[21] because the players are not a part of the same time, but in fact, have private and competitive utilities. The fact that Mechanism design optimizes the behaviour of competing players has led to broad applications spanning different fields including economics, politics, wireless networks, social networks, internet advertising, spectrum and bandwidth trading, logistics, supply chain, management, grid computing, and resource allocation problems in decentralized systems [22]–[28].

Advantages

(i) feasible,

(ii) budget balanced,

(iii) Individual rational, and

(iv) strongly implementable at the equilibria of the induced game.

We prove the existence of at least one generalized Nash equilibrium and show that or mechanism induces a Pareto efficient equilibrium.

4.SYSTEM ARCHITECTURE

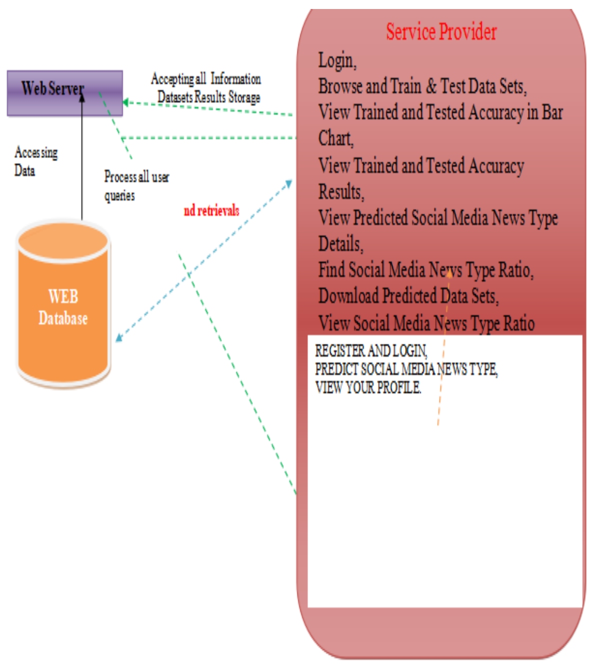

Figure.1 System Architecture

4.1 ALGORITHMS:

Decision tree classifiers

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, …, Ck is as follows:

Step 1. If all the objects in S belong to the same class, for example Ci, the decision tree for S consists of a leaf labeled with this class

Step 2. Otherwise, let T be some test with possible outcomes O1, O2,…, On. Each object in S has one outcome for T so the test partitions S into subsets S1, S2,… Sn where each object in Si has outcome Oi for T. T becomes the root of the decision tree and for each outcome Oi we build a subsidiary decision tree by invoking the same procedure recursively on the set Si.

Gradient boosting

Gradient boosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees.[1][2] When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest.A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

K-Nearest Neighbors (KNN)

- Simple, but a very powerful classification algorithm
- Classifies based on a similarity measure
- Non-parametric
- Lazy learning
- Does not "learn" until the test example is given
- Whenever we have a new data to classify, we find its K-nearest neighbors from the training data

Example

- Training dataset consists of k-closest examples in feature space
- Feature space means, space with categorization variables (non-metric variables)
- Learning based on instances, and thus also works lazily because instance close to the input vector for test or prediction may take time to occur in the training dataset

Logistic regression Classifiers

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name logistic regression is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name multinomial logistic regression is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar.

Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

Naïve Bayes

The naive bayes approach is a supervised learning method which is based on a simplistic hypothesis: it assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature .

Yet, despite this, it appears robust and efficient. Its performance is comparable to other supervised learning techniques. Various reasons have been advanced in the literature. In this tutorial, we highlight an explanation based on the representation bias. The naive bayes classifier is a linear classifier, as well as linear discriminant analysis, logistic regression or linear SVM (support vector machine). The difference lies on the method of estimating the parameters of the classifier (the learning bias).

While the Naive Bayes classifier is widely used in the research world, it is not widespread among practitioners which want to obtain usable results. On the one hand, the researchers found especially it is very easy to program and implement it, its parameters are easy to estimate, learning is very fast even on very large databases, its accuracy is reasonably good in comparison to the other approaches. On the other hand, the final users do not obtain a model easy to interpret and deploy, they does not understand the interest of such a technique.

Thus, we introduce in a new presentation of the results of the learning process. The classifier is easier to understand, and its deployment is also made easier. In the first part of this tutorial, we present some theoretical aspects of the naive bayes classifier. Then, we implement the approach on a dataset with Tanagra. We compare the obtained results (the parameters of the model) to those obtained with other linear approaches such as the logistic regression, the linear discriminant analysis and the linear SVM. We note that the results are highly consistent. This largely explains the good performance of the method in comparison to others. In the second part, we use various tools on the same dataset (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b and RapidMiner 4.6.0). We try above all to understand the obtained results.

Random Forest

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for decision trees' habit of overfitting to their training set.

Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance.

The first algorithm for random decision forests was created in 1995 by Tin Kam Ho[1] using the random subspace method, which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.

An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).The extension combines Breiman's "bagging" idea and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance.

Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SVM

In classification tasks a discriminant machine learning technique aims at finding, based on an *independent and identically distributed* (*iid*) training dataset, a discriminant function that can correctly predict labels for newly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminant classification function takes a data point *x* and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminant approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric perspective, learning a classifier is equivalent to finding the equation for a multidimensional surface that best separates the different classes in the feature space.

SVM is a discriminant technique, and, because it solves the convex optimization problem analytically, it always returns the same optimal hyperplane parameter—in contrast to *genetic algorithms* (*GAs*) or *perceptrons,* both of which are widely used for classification in machine learning. For perceptrons, solutions are highly dependent on the initialization and termination criteria. For a specific kernel that transforms the data from the input space to the feature space, training returns uniquely defined SVM model parameters for a given training set, whereas the perceptron and GA classifier models are different each time training is initialized. The aim of GAs and perceptrons is only to minimize error during training, which will translate into several hyperplanes' meeting this requirement.

5.MODULES

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Browse and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Predicted Social Media News Type Details, Find Social Media News Type Ratio, Download Predicted Data Sets, View Social Media News Type Ratio Results,, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN,PREDICT SOCIAL MEDIA NEWS TYPE, VIEW YOUR PROFILE.

6.CONCLUSION

Our primary goal in this paper was to design a mechanism to induce a GNE solution in the misinformation filtering game, where (i) each platform agrees to participate voluntarily, and (ii) the collective utility of the government and the platforms is maximized. We designed a mechanism and proved that it satisfies these properties along with budget balance. We also presented an extension of the mechanism with weaker technical assumptions.

Ongoing work focuses on improving the valuation and average trust functions of the social media platforms based on data. We also consider incorporating uncertainty in a platform's estimates of the impact of their filter. These refinements of the modeling framework will allow us to make our mechanism more practical for use in the real world.

Future research should include extending the results of this paper to a dynamic setting in which the social media platforms react in real-time to the proposed taxes/subsidies. In particular, someone could develop an algorithm that the players can use to iteratively arrive at the Nash equilibrium. In such an algorithm, the social planner can receive additional information from the players while they iteratively learn the GNE. Then, she can use this information to change her allocations dynamically, allowing us to relax either Assumption 5 on monitoring of average trust, or Assumption 6 on the excludability of the platforms.

7.REFERENCES

[1] W. Davies, “The age of post-truth politics,” The New York Times, vol. 24,p. 2016, 2016.

[2] J. Cone, K. Flaharty, and M. J. Ferguson, “Believability of evidence

matters for correcting social impressions," Proceedings of the NationalAcademy of Sciences, vol. 116, no. 20, pp. 9802–9807, 2019.

[3] Z. Tufekci, "Youtube, the great radicalizer," The New York Times, vol. 10,2018.

[4] A. D. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks,"Proceedings of the National Academy of Sciences, vol. 111, no. 24, pp. 8788–8790, 2014.

[5] J. Weedon, W. Nuland, and A. Stamos, "Information operationsand facebook," Retrieved from: https://fbnewsroomus. files. wordpress. com/2017/04/facebook-and-information-operations-v1. pdf, 2017.

[6] H. Farrell and B. Schneier, "Common-knowledge attacks on democracy,"Berkman Klein Center Research Publication, no. 2018-7, 2018.

[7] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," Journal of economic perspectives, vol. 31, no. 2, pp. 21136, 2017.

[8] O. Analytica, "Russia will deny cyberattacks despite more us evidence,"Emerald Expert Briefings, no. oxan-db, 2018.

[9] A. Bessi, M. Coletto, G. A. Davidescu, A. Scala, G. Caldarelli, and W. Quattrociocchi, "Science vs conspiracy: Collective narratives in the age of misinformation," PloS one, vol. 10, no. 2, p. e0118093, 2015.

[10] E. Brown, "Propaganda, misinformation, and the epistemic value of democracy," Critical Review, vol. 30(3–4), pp. 194–218, 2018.

[11] J. A. Tucker, Y. Theocharis, M. E. Roberts, and P. Barber´a, "From liberation to turmoil: Social media and democracy," Journal of Democracy, vol. 28(4), pp. 46–59, 2017.

[12] A. Sternisko, A. Cichocka, and J. J. Van Bavel, "The dark side of social movements: Social identity, non-conformity, and the lure of conspiracy theories," Current opinion in psychology, vol. 35, pp. 1–6, 2020.

[13] R. Jaakonm¨aki, O. M¨uller, and J. Vom Brocke, "The impact of content, context, and creator on user engagement in social media marketing,"Proceedings of the 50th Hawaii international conference on system sciences, 2017.

[14] O. Candogan and K. Drakopoulos, "Optimal signaling of content accuracy: Engagement vs. misinformation," Operations Research, vol. 68, no. 2, pp. 497–515, 2020.

[15] E. A. Vogels, A. Perrin, and M.

Anderson. (2020) Most americans think social media sites censor political viewpoints. [Online]. Available: https://www.pewresearch.org/internet/2020/08/19/most-americans-think-social-media-sites-

[16] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, 2018.

[17] A. Mas-Colell, M. D. Whinston, and J. R. Green, Microeconomic theory. Oxford University Press, 1995.

[18] A. Dave and A. Malikopoulos, "The prescription approach to decentralized stochastic control with word-of-mouth communication," arXive-prints, p. arXiv:1907.12125, Sep 2019.

[19] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Y¨uksel, "Informationstructures in optimal decentralized control," in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC). IEEE, 2012, pp. 1291–1306.

[20] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach,"IEEE Transactions on Automatic Control, vol. 58, no. 7, pp. 1644–1658,2013.

[21] A. A. Malikopoulos, C. G. Cassandras, and Y. J. Zhang, "A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections," Automatica, vol. 93, no. April, pp. 244–256,2018.

[22] S. Sharma and D. Teneketzis, "Local public good provisioning in networks: A Nash implementation mechanism," IEEE Journal on Selected Areas in Communications, vol. 30, no. 11, pp. 2105–2116, 2012.

[23] A. Sinha and A. Anastasopoulos, "Generalized proportional allocation mechanism design for multi-rate multicast service on the internet," 51st Annual Allerton Conference on Communication, Control, andComputing (Allerton), pp. 146–153, 2013.

[24] A. Kakhbod and D. Teneketzis, "An efficient game form for unicast service provisioning," IEEE Transactions on Automatic Control, vol. 57,no. 2, pp. 392–404, 2011.

[25] R. Jain and J. Walrand, "An efficient nash-implementation mechanism for network resource allocation," Automatica, vol. 46(8), pp. 1276–1283,2010.