Abstract—The concept of “truth,” as a public good is the production of a collective understanding, which emerges from a complex network of social interactions. The recent impact of social networks on shaping the perception of truth in political arena shows how such perception is corroborated and established by the online users, collectively. However, investigative journalism for discovering truth is a costly option, given the vast spectrum of online information. In some cases, both journalist and online users choose not to investigate the authenticity of the news they receive, because they assume other actors of the network had carried the cost of validation. Therefore, the new phenomenon of “fake news” has emerged within the context of social networks. The online social networks, similarly to System of Systems, cause emergent properties, which makes authentication processes difficult, given availability of multiple sources. In this study, we show how this conflict can be modeled as a volunteer’s dilemma. We also show how the public contribution through news subscription (shared rewards) can impact the dominance of truth over fake news in the network.

Keywords—disseminating truth, volunteer’s dilemma; fake news; social network policy; incentivizing mechanism; cooperative governance

I. INTRODUCTION

In today’s social network era, the perception of truth depends on emerging social interactions along with the traditional news broadcasters. These online social interactions leaves us with a diverse pool of news resources most of which are all but impossible to be individually fact checked and verified in terms of credibility and truthfullness. The lightening spread of news in the social media is known to be one aspect of network diffusion in social networks [1]. The network effect in disseminating news suggests that a piece of news gains credibility when goes viral. In other words, higher number of redistribution for a piece of news by other nodes is often perceived as its authenticity across a networked system [2] [3] [4] [5]. Activity of these nodes (agents) also depends on other agents in complex network of social interactions. At the same time, accessing the stream of true and credible news is costly. However, creating such credible streams in social networks to eliminate dominancy of fake news will be ultimately beneficial for the whole society, in long run. This suggests that the truthfull news (i.e. truth) can be considered as a public good.

In networked social systems with public goods, public service has cost for volunteering agents while it generates benefit that belongs to all members [6]. In social networks, disseminating truth is considered a public good as costly process of attaining the truth and the public benefit of the news disintencivize the social members to volunteer in the public process and instead incentivize them to rely on other agents in the networked system. Public good game in social systems is a model of situation where the presence of free-rider incentives jeopardizes the socially beneficial outcome [7]. Accordingly, in the volunteer’s dilemma in social systems, higher number of social members results in lower level of active participation in validating and authenticating the news (public goods). This conflict towards the volunteering equilibrium is raised in many forms: cooperation among the members [8] [9], inclusion of heterogeneous agents among the players [10], which is often referred to as “strong” and “weak” members [9] [11], punishing defectors in the system [7] [12] [13], rewarding the volunteers using shared reward pool [14] [15], and so forth. The incentivising mechanism for such circumstances is among the well studied areas and multiple studies have shown the mathematical merits of it, including the existence and stability of its equilibrium points in evolutionary games [13] [16]. The incentivizing mechanism and the principle dissemination, as two basic ideas behind this study, are two pillars of governance in systems of systems [17].

Game theoretical approaches are applied to model the equilibrium in social systems with heterogeneous groups of agents. The attacker-defender game is applied to protect power system against intelligent external attacks [18]. A two-step game-based formulation can optimize the cloud service request by distributed mobile devices [19]. A game theoretic payment sharing model could incentivize the truthfull participation of agents in energy market [20] [21]. Finally, a mathematical framework that incentivizes the agent movements also reduces the collective cost of a transportation network [22].
In this study, we introduce a mathematical integrated framework, based on game theoretic mixed-strategy formulation of volunteer’s dilemma, to model disemination of truth and fake news in social networks.

A. Basic Model Assumptions

In this study we introduce two types of agents within the population: regular agents and the fake news agents. We assume the former population is greater than the latter in size (10-fold) when the collective benefit of disemination of truth (public good) belongs to all players and the group benefit of disemination of fake news merely belongs to the group of fake news agents. Also, truth is assumed to be outcome of social intereaction and dominance between the regular and fake news groups. We assume that the agents are distributed and network is balanced among agents (either strategically or randomly), which implies that the group with higher number of volunteers can expect the dominance of its message in the social network. For rewardign mechanism, we assume that the subscription based journalism works as a rewarding mechanism and volunteering agents receive more or less the same share of the reward. We also assume that the effects of fake news is not directly seen by the regular agents, i.e. regular agents cannot categorically and directly punish fake news agents. An implication of the latter assumption is that the equilibrium of fake news model doesn’t affect the equilibrium of regular news. The opposite statement is not true. Accordingly, we formulate the probabilistic model of fake news based on the equilibrium of probabilitic model of regular users. Finally, we assume that the users among each group of agents (regular and fake news agents) are homogenous, then the model is symetric for each group.

In the following section, the volunteer’s dilemma is illustrated, then the modified formulation of symmetric volunteer’s dilemma (VOD) is expressed, and the payoff formulations of fake news agents are presented. In the third section, we show the analytical equilibrium of volunteering model defined as the net payoff versus the volunteering ratio. The effect of shared reward and group size is shown on the equilibrium points. Lastly, we calculate the equilibrium points for diseminating the fake news based on the model equilibrium of regular users.

II. MODEL

To start with, we introduce the model variables and parameters. A number of regular agents (N) are interacting in social system. These regular agents can either participate (M) in validating and authenticating the truthful news or consume news (N-M). The minimum group size of volunteers is required to achieve the public good (k). The cost of volunteering (c) and the cost of failing to achieve the public good (α) occurs in case. Having an rewarding mechanism, σ is distributed among the volunteering agents. Because the model is a symetric model, the equilibrium shows the volunteering ratio or the probability of volunteering among users (p). On the other side of aisle, F number of fake news agents with cost $c_f$ volunteer in publishing fake news. Both are fraction of N and c respectively. In the rest of this paper, we refer to act of not volunteering (sharing/distributing but not authenitcating and validating) as defecting. We formulate the effective features for agent behavior in social networks and express the model equations next. Table 1 shows the variables, mathematical symbols, and their description.

| Variable | Table column subhead |
|----------|----------------------|
| N        | Number of regular users that could potentially disseminate truth in the network |
| M        | Number of volunteering agents among regular agents |
| c        | Cost of volunteering for regular agents |
| α        | Cost of failing to dominate the news cycle |
| k        | Minimum group size to dominate the news cycle: the number of volunteering fake news agents |
| x        | Volunteering ratio and the probability of volunteering in symetric model |
| σ        | Aggregated reward shared among regular agents |
| $p_v(M)$ / $p_d(k)$ | Payoff of volunteering for a regular/fake news agent as function of the number of volunteers |
| $p_v(M)$ / $p_d(k)$ | Payoff of defecting for a regular/fake news agent as function of the number of volunteers |
| $F_v(k)$ / $F_d(k)$ | Average payoff of volunteering/defecting regular agents as function of group size |
| F        | Number of fake news users that could potentially disseminate fake news |
| $c_f$    | Cost of volunteering for fake news agents |
| $F_v(M)$ / $F_d(M)$ | Average payoff of volunteering/defecting fake news agents as a function of regular volunteering agents |
| $p^*$    | Equilibrium probability of volunteering for regular agents |
| $p_v$ / $p_d$ | Expected value of volunteering/defecting payoff of fake news agents |

A. Volunteer’s Dilemma

Equation 1 shows the individual payoff in volunteer’s dilemma (VOD) for N players. At each round, N random social network users, among infinite number of people, and potential volunteers among them are selected. The players produce the public good if at least k players decide to volunteer (1≤k≤N). The cost of volunteering is c, relative the baseline payoff of 1, and the cost is mandatory whether or not the public good is achieved. If the number of volunteers is less than k, the failing cost of $α>c$ occurs for each player.

\[
p_v(M) = \begin{cases} 1-c & \text{if } M \geq k \\ 1 - a - c & \text{if } M < k \end{cases} \quad p_d(M) = \begin{cases} 1 & \text{if } M \geq k \\ 1 - a & \text{if } M < k \end{cases}
\]

(1)

where M is number of volunteers and k is minimum number of volunteers to achieve public good.

B. Payoff Model of Truth Disemmination

In the general VOD with shared reward among the volunteers, the average payoff for volunteering is formulated by modifying the reward sharing model of generalized volunteer’s dilemmas (Chen, 2013). To express the equations, we unify the terminology of variables in the model with the
introduced variables in Table 1. Accordingly, number of players represents regular users and number of volunteers represents users volunteering in disseminating truth, etc. Eq. 2 shows the average payoff of volunteering agents having a fixed shared reward distributed among them:

\[
\overline{P}_v(k) = \sum_{M=k}^{N-1} \left( \begin{array}{c} N-1 \\ M \end{array} \right) x^M(1-x)^{N-M-1}(1-c) + \left[ 1 - \sum_{M=k}^{N-1} \left( \begin{array}{c} N-1 \\ M \end{array} \right) x^M(1-x)^{N-M-1}(1-c-a) \right]
\]

(2).

The average payoff of regular users when they don’t participate in disseminating the truthful news is expressed as:

\[
\overline{P}_d(k) = \sum_{M=k}^{N-1} \left( \begin{array}{c} N-1 \\ M \end{array} \right) x^M(1-x)^{N-M-1} + \left[ 1 - \sum_{M=k}^{N-1} \left( \begin{array}{c} N-1 \\ M \end{array} \right) x^M(1-x)^{N-M-1}(1-a) \right] - \sigma \overline{P}_v
\]

(3)

where \( \sigma \) is shared reward among volunteers and \( x \) is the ratio of volunteers or volunteering probability.

C. Payoff Model of Fake News Dissemination

In the previous section, we modified a model of VOD for regular social media users. We also introduce another model of VOD with different formulation for users who are incentivized to disseminate fake news. While payoff mechanism for the former group of users is formulated in Eq. 1, the payoff model of latter groups of users include different parameters and excludes the rewarding mechanism. Here, the intuition is the fake news agents cannot benefit from dominance of truth in social network and don’t contribute to dissemination of truth through shared rewardign mechanism. Eq. 4 shows the individual payoff of fake news agents and Eq. 5 and 6 show the average payoff of volunteering and defecting to contribute to fake news community respectively:

\[
p_v(k) = \begin{cases} 
1-cf & \text{if } k \geq M \\
1-a-cf & \text{if } k < M
\end{cases}
\]

\[
p_d(k) = \begin{cases} 
1 & \text{if } k \geq M \\
1-a & \text{if } k < M
\end{cases}
\]

(4)

\[
\overline{P}_v(M) = \sum_{k=M-1}^{F-1} \left( \begin{array}{c} F-1 \\ k \end{array} \right) x^k(1-x)^{F-k-1}(1-cf)
\]

\[
+ \left[ 1 - \sum_{k=M-1}^{F-1} \left( \begin{array}{c} F-1 \\ k \end{array} \right) x^k(1-x)^{F-k-1}(1-cf-a) \right] - \sigma \overline{P}_v(M)
\]

(5)

D. Integrated Incentivizing Model

In social network, we combine the potentially truthful and fake news agents. The former set of users share the collective reward of disseminating truth in the network. The second set of agents could achieve in-group benefits of spreading fake news when disseminating truth affects their utility through the cost of failure. In the combination of these agents in network, the smaller and less public group, i.e. fake news agents, calculate their expected payoff as a function of expected number of volunteers among regular agents. The volunteering ratio being estimated from the Eq. 2 and 3 equilibrium, fake news agents infer a binomial distribution for the number of volunteering agents. Eq. 7 shows the probabilistic sum of averaged payoff for fake news agents using probability mass function (pmf) of volunteering agents:

\[
\overline{P}_v = \sum_{M=0}^{F-1} \left( \begin{array}{c} N \\ M \end{array} \right) p^M(1-p^*)^{F-M} \overline{P}_v(M)
\]

(6)

(7)

where \( v \) represents the volunteering and \( d \) represents the defecting group. This expected payoff shows the potential incentive for fake news agents to actively participate in disseminating fake news in the social network. In effect, this model shows the effects of rewarding mechanism on incentivizing the dominance of truth in the network. In the next section, we explore the analytical effect of rewarding mechanism and model parameters on symmetric equilibrium points for disseminating truthful and fake news.

III. RESULTS

To show the analytical results, we first review our numeric assumptions for parameters of the integrated model. We assume one hundred regular users in the social network. Around one twelfth of this number are assumed as potential fake news agents (8). We create a fictional currency and represent it by \( \mathcal{C} \). We pick \( \mathcal{C}0.5 \) for volunteering cost (c) and \( \mathcal{C}0.9 \) for failing cost (a). As it was mentioned before, the cost of creating fake news is a fraction of cost of volunteering for disseminating truth (cf = \( \mathcal{C}0.1 \)). For the rewarding mechanism, we assume a shared reward between \( \mathcal{C}5 \) and \( \mathcal{C}8 \), which is evenly distributed among the volunteering agents. For the minimum volunteering group size, based on the maximum number of active fake news sources, we assume a range of numbers less than 8.
To run the model and show the results, we define two comparing factors (Table 2). The first factor calculates the difference between averaged payoff of volunteering and defecting agents. Apparently, the measure shows whether the agents benefit from either volunteering or defecting depending on the difference being positive or negative. For the second factor, we consider the fake news agents and based on Eq. 2 and 3, we calculate the expected difference between the payoff of volunteering and defecting (Eq. 7) among fake news agents.

| Factor | Description |
|--------|-------------|
| \( P_v(k) - P_d(k) \) | Net difference between the average payoff of volunteering and defecting (regular agents). Net value equal to zero shows a potential equilibrium for volunteering ratio among regular agents. (Fig. 1 and 2) |
| \( \bar{P}_v - \bar{P}_d \) | Net difference between the expected payoff of volunteering and defecting among fake news agents, which shows whether the “fake news” agents are incentivized to volunteer or defect. Net value equal to zero indicates a potential equilibrium for volunteering ratio among the fake news agents. (Fig. 3) |

To explore and discuss the analytical results, first we show the possible equilibrium points for volunteering ratio among the regular agents. We approach the equilibrium points from two perspective: the effect of shared reward on the expected payoff of volunteering and the that of minimum group size on both metrics. Then, we discuss the effect of model equilibriums on the expected payoff of volunteering among fake news agents.

Increasing the shared reward among the volunteering agents significantly affects the volunteering payoff and the potential equilibrium points. Fig. 1 shows the net payoff corresponding to the volunteering probability for a range of total shared rewards. In this case, the group size is selected equal to 6, which represents minimum number of volunteering agents to achieve the public good. In this figure, the net payoff equal to zero creates two sets of equilibrium points. Both sets show equilibrium because for the net payoff of zero, agents don’t have incentive to either increase or decrease their volunteering probability. However, one set of those equilibrium points are stable while the other set are unstable points. Intuitively, without proving the stability by the mathematical expressions, the set with positive derivative is unstable because any change in volunteering ratio reinforces further change to the same direction (i.e. increasing ratio results in higher payoff for volunteering that incentivize higher ratio). Hence, the second set (with negative derivative) shows the stable equilibrium points for volunteering ratio. This figure shows that higher amount of shared reward leads to higher stable equilibrium points in the symmetric model. We know that the dominant strategy for low values of shared reward \( (\sigma<4) \) is defecting (Eq. 3 > Eq. 2).

In the second case, we show the effect of minimum group size on net payoff and equilibrium points. We select the range of 5 to 8 for group size (lower than the number of potential fake news agents) and 5 for shared reward. Fig 2 shows that the group size has considerable effect on the first set of equilibrium points but negligible effect on the second set (stable). A reason is low variance of binomial distribution because of relatively high number of regular agents. This means that certain probability either pass the group threshold or doesn’t and it is very improbable to fall in between. Accordingly, the higher the group size is, the lower the first equilibrium will be. For high number of group size (\( >8 \)), the dominant individual strategy is defecting.

In our model, dissemination of fake news happens when the number of fake news sources dominates the number of truthful sources \( (k>M) \) in Eq. 5 and 6. The fake news agents might be incentivized to volunteer knowing about the volunteering preference among regular users. In this study, we claim that the rewarding mechanism for regular users could be designed as disincetivizing mechanism for participation among fake news agents. Lower volunteering among fake news agents means higher probability of dissemination of truth in the network. Having these being said, we implement the Eq. 7 using the equilibrium points for regular users \( (p* \) in Eq. 2 and 3). In Fig. 3, we show the net payoff for volunteering and defecting among the fake news agents. The selected range of volunteering equilibrium points (legends) shows the realistic points calculated using the first model (Fig. 1 and 2). The expected volunteering ratio has significant effect on the fake news equilibrium points. For lower volunteering points, the net payoff of participation in disseminating fake news increases and the equilibrium points shift toward the lower values, which means that relatively lower level of activity among fake news agent can incentivize those agents to remain active. Another effect of volunteering ratio is seen on the amount of net payoff for participation among the fake news agents. The potential net payoff is significantly higher when the volunteering ratio is expected to be lower, i.e. 4 percent volunteering probability shows maximum net payoff 4-fold of the case with 10 percent probability. Finally, higher volunteering ratio shifts the first set of equilibrium points (unstable) to right. Although these points are unstable, its implication on adopted strategy by fake news agents is notable: any participation rate lower than those points could incentivize lowering the participation rate to zero.

![Figure 1. Net payoff difference for volunteering and defecting agents corresponding to users’ volunteering ratio. The \( \sigma \) shows total reward (cost) distributed (shared) among volunteers (users).](image-url)
negative derivative of net payoff).

Last but not least, the reverse effect of equilibrium points among fake news agents on the main model should be discussed. The participation rate among the fake news agents affects the volunteering model of regular users through group size (k). Increasing the number of participants among fake news agents also increases the minimum group size to achieve the public good by the regular users. However, as was discussed before, the effect of the group size on the stable equilibrium points in Fig. 2 is minimal, i.e. the equilibrium point of volunteering remains around 0.09 for range of feasible group size (<8). This imply that the regular users are minimally incentivized by equilibrium behavior of fake news agents.

IV. DISCUSSION AND CONCLUSION

In sum, the effect of real-world financial behavior among social network users, such as subscription to credible journalism, on dissemination of fake news can be discussed using analytical game theoretic models. In this paper, we proposed an integrated model combining two VOD models of volunteering among regular users and fake news agents in social network. The integrated model attacks the systemic problem of emerging fake news against the fact-based news (truth) in social networks, which manifests a complex System of Systems. To show the impact of agent adopted policies on spreading fake news in networked social systems, we formulated a rewarding mechanism that incentivizes volunteering in the system toward dissemination of truth. The equilibrium points represent the agent adopted policy in a symetric system of systems. This particular mechanism (agent adopted policy) uses a shared subscription program with a fixed subscription fee for regular agents (users). We formulated a payoff mechanism for fake news agents. Integrating the volunteering model with the fake news participation model, we observed the effects of a shared rewarding mechanism on the equilibrium points of activity among fake news agents. We showed that relatively marginal variation in shared reward has significant effect on the dissemination and ultimate dominance of fake news in the network. We discussed that the equilibrium points for fake news agents heavily depends on the incentivized system equilibrium while the reverse effect is not observable.

For future studies, the mathematical illustration of the integrated equilibriums helps to establish a mechanism which incentivizes the dominance of truth in social networks. In addition, considering asymmetric models with heterogeneous agents can considerably improve the explanatory aspect of our proposed model. Finally, the dynamic algorithms can be applicable to design rewarding mechanisms that target collective benefit of the System of Systems.

REFERENCES

[1] Y. a. J. J. Jiang, "Diffusion in social networks: A multiagent perspective," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 45, no. 2, pp. 198--213, 2015.

[2] S. M. a. Z. X. a. S. M. Shariff, "User perception of information credibility of news on twitter," in European Conference on Information Retrieval, 2014.

[3] G. a. H. C. Szab, "Phase transitions and volunteering in spatial public goods games," Physical review letters, vol. 89, no. 11, p. 118101, 2002.

[4] A. a. P. M. Szolnoki, "Antisocial pool rewarding does not deter public cooperation," in Proc. R. Soc. B, vol. 282, The Royal Society, p. 20151975.

[5] G. a. G. T. a. D. R. a. K. K. a. B. R. a. Iniguez, "Effects of deception in social networks," in Proc. R. Soc. B, 2014.

[6] R. M. a. W. J. M. Isaac, "Group size effects in public goods provision: The voluntary contributions mechanism," The Quarterly Journal of Economics, vol. 103, no. 1, pp. 179--199, 1988.

[7] B. C. T. a. S. G. Herrmann, "Antisocial punishment across societies," Science, vol. 319, no. 5868, pp. 1362-1367, 2008.
[8] C. D. M. S. H. J. &. S. K. Hauert, "Volunteering as red queen mechanism for cooperation in public goods games," *Science*, vol. 296, no. 5570, pp. 1129-1132, 2002.

[9] A. Diekmann, "Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence," *International Journal of Game Theory*, vol. 22, no. 1, pp. 75-85, 1993.

[10] F. C. a. S. M. D. a. P. J. M. Santos, "Social diversity promotes the emergence of cooperation in public goods games," *Nature*, vol. 454, no. 7201, pp. 213-216, 2008.

[11] J.-Z. a. W. R.-W. a. L. Y.-T. He, "Evolutionary stability in the asymmetric volunteer's dilemma," *PloS one*, vol. 9, p. e103931, 2014.

[12] D. a. S. A. a. P. M. a. S. G. Helbing, "Punish, but not too hard: how costly punishment spreads in the spatial public goods game," *New Journal of Physics*, vol. 12, no. 8, p. 083005, 2010.

[13] A. a. P. W. Diekmann, "Punitive preferences, monetary incentives and tacit coordination in the punishment of defectors promote cooperation in humans," *Scientific reports*, vol. 5, p. 10321, 2015.

[14] X. a. G. T. a. D. U. Chen, "Shared rewarding overcomes defection traps in generalized volunteer's dilemmas," *Journal of theoretical biology*, vol. 335, pp. 13-21, 2013.

[15] T. a. U. S. Sasaki, "Rewards and the evolution of cooperation in public good games," *Biology letters*, vol. 10, no. 1, p. 20130903, 2014.

[16] Z. a. K. S. a. J. M. a. T. J. Wang, "Universal scaling for the dilemma strength in evolutionary games," *Physics of life reviews*, vol. 14, pp. 1--30, 2015.

[17] H. R. A. G. a. M. Darabi, "Governance mechanism pillars for systems of systems," in *System of Systems Engineering (SoSE)*, Genova, 2012.

[18] L. M. A. H. S. A. &. S. A. I. Wei, "Defending mechanisms for protecting power systems against intelligent attacks," in *System of Systems Engineering Conference (SoSE)*, San Antonio, 2015.

[19] Y. L. X. &. P. M. Wang, "A nested two stage game-based optimization framework in mobile cloud computing system," in *Service Oriented System Engineering (SOSE)*, Genova, 2013.

[20] A. a. B. H. Ehsanfar, "Interactive Multi-Consumer Power Cooperatives with Learning and Axiomatic Cost and Risk Disaggregation," in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, 2015.

[21] A. a. B. H. Ehsanfar, "An Incentive-Compatible Scheme for Electricity Cooperatives: An Axiomatic Approach," *IEEE Transactions on Smart Grid*, 2016.

[22] A. S. F. a. M. T. I. Ehsanfar, "An asymmetric intelligent model for public transportation networks," in *11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, 2008.

[23] J. Pena, "GROUP-SIZE DIVERSITY IN PUBLIC GOODS GAMES," *Evolution*, vol. 66, pp. 623--636, 2012.

[24] M. a. A. K. a. W. A. Rafalak, " Incredible: Is (Almost) All Web Content Trustworthy? Analysis of psychological factors related to website credibility evaluation," in Rafalak, Maria and Abramczuk, Katarzyna and Wierzbicki, Adam, 2014.

[25] C. a. M. M. a. P. B. Castillo, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*, 2011.