scientific reports

Check for updates

OPEN

Protecting infrastructure performance from disinformation attacks

Saeed Jamalzadeh[1], Kash Barker[1]✉, Andrés D. González[1] & Sridhar Radhakrishnan[2]

Disinformation campaigns are prevalent, affecting vaccination coverage, creating uncertainty in election results, and causing supply chain disruptions, among others. Unfortunately, the problems of misinformation and disinformation are exacerbated due to the wide availability of online platforms and social networks. Naturally, these emerging disinformation networks could lead users to engage with critical infrastructure systems in harmful ways, leading to broader adverse impacts. One such example involves the spread of false pricing information, which causes drastic and sudden changes in user commodity consumption behavior, leading to shortages. Given this, it is critical to address the following related questions: (i) How can we monitor the evolution of disinformation dissemination and its projected impacts on commodity consumption? (ii) What effects do the mitigation efforts of human intermediaries have on the performance of the infrastructure network subject to disinformation campaigns? (iii) How can we manage infrastructure network operations and counter disinformation in concert to avoid shortages and satisfy user demands? To answer these questions, we develop a hybrid approach that integrates an epidemiological model of disinformation spread (based on a susceptible-infectious-recovered model, or SIR) with an efficient mixed-integer programming optimization model for infrastructure network performance. The goal of the optimization model is to determine the best protection and response actions against disinformation to minimize the general shortage of commodities at different nodes over time. The proposed model is illustrated with a case study involving a subset of the western US interconnection grid located in Los Angeles County in California.

Well-publicized disinformation campaigns surrounding recent US Presidential elections and the adoption of pandemic-related vaccinations have increased awareness among researchers that historical problems of misinformation / disinformation are exacerbated due to the wide availability and use of online platforms. Disinformation, defined as information that falsely characterizes the state of the system, including rumors, factual errors, and attempts at deception[1], is rising on online platforms[2,3].

There is substantial literature on modeling the effects of and protection against false data injections by adversaries and connections to the operability and functionality of critical infrastructures[4–6]. However, an over-the-horizon problem may result from an adversary that seeks to attack critical infrastructure indirectly by altering the consumption behavior of human intermediaries who are influenced by weaponized disinformation distributed by the adversary.

Consider the following plausible scenarios that are extended from collections of actual events. An airline passenger could tweet an alert about a suspicious package, which, if shared rapidly and widely, could cause significant delays in flights and major traffic jams on many primary, secondary, and tertiary roads (similar to what was experienced at London's Gatwick airport[7]). Hackers could compromise a major US pipeline network, but the rampant spread of misinformation leads to a dramatic escalation in the aftermath and a physical, real-world increase in gas prices (similar to what was experienced when a news network spread a false story about a Russian hack of the US power grid[8]). Finally, false reports of accidents on social media could lead to dynamic rerouting of drivers, causing congestion in particular areas subject to attack (similar events have occurred globally[9–11] and could worsen with the emergence of autonomous vehicles[12]).

To begin to address some of these scenarios, we develop a model to examine the interactions between information/disinformation spread, subsequent commodity consumption behavior, and the resulting infrastructure network balance. To do so, we integrate an epidemiological model of information/disinformation spread with

[1]School of Industrial and Systems Engineering, University of Oklahoma, Norman, OK 73019, USA. [2]School of Computer Science, University of Oklahoma, Norman, OK 73019, USA. ✉email: kashbarker@ou.edu

a network flow model. We relate the two models through human intermediaries who adopt information/disinformation that changes the way that they interact with the infrastructure network.

Contributions of our integrated model include: (i) we account for the evolution of disinformation spread over time based on the outcomes of virtual interaction between pair of users in social media, ultimately projecting user behavior onto commodity consumption, (ii) we introduce an information protection mechanism to combat against disinformation spread, and (iii) we develop a mixed-integer programming formulation to balance the performance of the critical infrastructure network and plan for targeting (good) information to counter disinformation.

This paper is organized as follows. The Background and Literature section provides a methodological background on the concepts of disinformation spread and critical infrastructure network optimization along with the associated literature. The subsequent model section explains the proposed integration of epidemiological and mathematical programming models. The Case Study section illustrates the proposed model with a case study involving the power distribution network in Los Angeles County, California. Finally, the Conclusion section offers concluding remarks and future research opportunities.

Background and literature review

Our proposed work relies on two key areas: (i) spread of information and disinformation, and (ii) network flow models for infrastructure. In this section, we offer a review of these two areas, detailing some of the current research gaps addressed in this paper.

Models of information and disinformation spread. Within social networks, people exchange information to ultimately influence others, where influence is defined as an action "to induce a change in the behavior of another that is in accordance with the wishes of the influencer"[13,14]. Each individual communicates with (or influences) many other peers and, similarly, individuals are influenced by numerous other peers. This influence can take on negative forms, such as information pollution, fake news, propaganda, misinformation, misinformation, disinformation, and hoaxes[15–17]. Social media users can become a source of online broadcast activity that affects personal and social behavior. Users may help speed up the transfer rate of information or disinformation and manipulate the content to match their points of view, deliberately or inadvertently, which may not be necessarily verified or verifiable. In such an online environment, if users do not pay enough attention to verified content and reliable sources of information, the information they receive may have varying levels of correctness and malicious intent.

The community of users can be classified into different categories based on how they respond to the influence of others. An analogy to the spread of influence and different categories of response is the spread of disease and different states of infection found in epidemiology literature[18]. A basic model for the spread of disease is the susceptible-infected-recovered (SIR) model, which uses a series of differential equations to describe the membership of different states at a point in time: those who are susceptible to the disease, those who are infected by it, and those who have recovered from it. An analogy can be made for those users reacting to information and disinformation. For example, for a group of power utility users who receive a fake message promoting a discount price for power usage during a specific time, those users may potentially share it with others or not, based on characteristics (e.g., personal traits) they exhibit. Using the SIR convention, individuals who adopt this disinformation and react to it directly by consuming more power can be classified as "Infected." Alternatively, users who are not influenced by this disinformation, for any reason, can be classified as "Removed." And users who have not received notification yet can be classified as "Susceptible." This classification of categories allows us to model, quantify, and predict their power usage during disinformation dissemination. There is a rich literature that formulates the phenomenon of transition between categorical labels with SIR models that employ a system of differential equations based on mean field theory or agent-based models that allow us to simulate the transmission of disinformation among autonomous agents in a flexible microscale manner[19–21].

Social media users are not limited to categories S, I, and R. For example, some groups of users intend to spread accurate information to fight against disinformation, or those who have already received disinformation but do not reshare it with other users, or those who have received disinformation but temporarily do not share it[22–24]. There may be communities on social media that spread authenticated information to counteract disinformation[22,25]. Furthermore, users in each category (that is, S, I, and R) can be classified as *aware* and *unaware*, where it is assumed that unaware users can become aware users based on contact with aware individuals at a given rate. Still, it is assumed that the reverse transition will not occur. We define the terminology "awareness" as knowledge and understanding that something is happening or exists[26]. In addition to the novel categories attached to classic SIR models to represent a community, some methods are developed to avoid bias originating from discretizing the solutions of SIR models[27].

Several different derivatives of the SIR class of models have been developed to extend the various categories of adoption of influence (e.g., information and disinformation)[28]. Given the link between networks and the spread of diseases[29,30], the SIR modeling enterprise has applications in other network-related applications: the spread of ideas[31–34] and the influence of social networks[35–39]. A related idea by[40] uses a variation of the SIR model to address the stifling of rumors. Still, it does not adequately allow for the competitive nature to describe the spread of information versus disinformation. It is because disinformation spreads differently than information, as noted in[41], with the former spreading faster and covering a large population on Twitter[2]. Social responses to disinformation will be examined by observing (i) how people evaluate information, (ii) how varying situations affect people's ability to evaluate information effectively, and (iii) how people act on information, including redistributing disinformation. In our context, *S* refers to individuals who have not yet been exposed to the disinformation content, *I* represents individuals who have heard the disinformation and changed their consumption behavior

as a result, and *R* represents individuals who have heard the disinformation but ignored them after realizing that the information they received was not true or accurate.

Models of infrastructure flow optimization. Complex infrastructure systems such as water, gas, transportation, and electricity are crucial for society's well-being and for promoting economic productivity. If one component of the system is affected by failure, larger spread effects can be experienced in other networks of infrastructures and networks of community members that suffer from unmet demand for goods and services[42]. As such, the resilience of critical infrastructure networks attracted researchers to study the ability of systems to mitigate the magnitude and duration of the components of the out-of-service infrastructure network[43,44].

Flow balance models are developed to determine how commodities are delivered from suppliers to customers, so that performance metrics such as average unsupplied demand and transportation costs are minimized, while guaranteeing that key operational constraints are observed. By the term "commodity," we broadly refer to flows of demanded entities (e.g., electric power, water, vehicles, data, goods) transmitted from one node to another through links connecting them. Several different flow balance optimization models are proposed in the literature that are applicable to infrastructures focused on disruptive events[45].

In the literature, there are numerous representations of network optimization of infrastructure networks. Hsu et al.[46] presented a generalized network flow model to model the long-term supply and demand of water resources. Tahiri et al.[47] proposed a network flow optimization model for similar water distribution networks, minimizing the total cost of meeting the demand for water. Martin et al.[48] optimized a gas network consisting of a set of compressors and pipes that connect the valves in order to minimize the total cost of the network subject to supply-demand balance. Banda et al.[49] similarly proposed a gas pipeline network optimization model that accounted for nonlinear isothermal equations. Traffic flow optimization problems have also been proposed[50]. Darayi et al.[51] proposed a multicommodity network flow optimization model to understand the criticality of different multimodal transportation nodes and links.

Especially important to the case study addressed subsequently are network optimization problems designed for electric power networks. Vasin et al.[52] proposed a model to optimize the flow of energy resources through a transportation network. Costa et al.[53] developed a two-stage linear programming model to reinforce power grids against attacks on transmission lines, proposing an exact algorithm to solve the model. Leuthold et al.[54] developed a nonlinear mixed-integer programming model to design an electricity market such that public welfare is maximized, with an application to the European electricity market. Wirtz et al.[55] proposed a sustainable multicommodity system design model with the power grid attached to the system using mixed integer linear programming. Electric power networks are critical sources of energy that enable the function of other infrastructures, and developing flow balance optimization models for electric power grids has become important for researchers[56], as have several network flow optimization representations of interdependent infrastructure networks that include electric power[57–60].

In this paper, we address the challenge of how to track and respond to disinformation attacks that disrupt infrastructure networks. Embedding the evolution of disinformation diffusion intensity over time attached to a flow balance optimization model of infrastructure network has two main benefits: (i) we can monitor and analyze the performance of infrastructure network disrupted by disinformation attacks over time, and (ii) act in opposition to disinformation propagation to mitigate the effect of disruption on the infrastructure network performance. To the best of our knowledge, such a model has not been proposed in the literature yet. To address this gap, we propose a network flow balance optimization model integrated with disinformation diffusion model that enables us to take opposite actions using social media to handle interruptions in infrastructure networks caused by disinformation attack.

Proposed integrated epidemiological + optimization (EPO) model

We propose and integrate two models to examine the interactive relationship between disinformation dissemination and critical infrastructure network performance: (i) the SIR model and (ii) a network flow balance optimization model. The network balance optimization model is used to balance critical infrastructure systems with respect to disinformation propagating on social networks, as the spread of disinformation on social networks affects the consumption behavior of social network users. These two networks are integrated in a multi-to-one environment from the social network to the critical infrastructure network, where communities of users are assigned to the set of infrastructure nodes.

SIR model. We describe the disinformation propagation process in the type of modeling "compartmental models" in which the population of social media users is divided into exclusive compartments. In such a formulation, we assign the rates at which the population within one compartment is transferred to another. In general, we can classify users into three exclusive compartments over time: (i) *S*, the proportion of users who are unaware of the disinformation and would have acted on it if known, (ii) *I*, the proportion of users who consumed the disinformation and changed (acted on the disinformation) their commodity usage schedule, and (iii) *R*, the proportion of users who were exposed to the disinformation but either ignored or detected it and are not interested in sharing it. Dividing the population of each compartment, we can formulate the dynamics of the compartments by replacing the size of the population with the proportion of the population.

The rate of transfer from one state to another is expressed as derivatives of the proportion of population in terms of time, and we make some assumptions to express the terms of the model. As such, we have a system of differential equations that describe how the proportion of people changes across different states over time by frequent communication. For example, given a population size *N*, for an unaware user randomly communicating with other users, the probability that the unaware user meets a user who adopted disinformation can be

expressed by $\frac{I}{N}$, and the rate of contact can be described as a coefficient of the total population, βN. Assuming that meetings between unaware users, S, and users who adopted disinformation result in the unaware user adopting disinformation, the population of unaware users decreases by βSI, and the population of users who adopted disinformation increases by the same size in time slots. Through this transformation process, users who consume disinformation have the opportunity to detect or ignore disinformation at a rate γ, with the associated population γI, and are no longer classified as the group that already consumes disinformation. Thus, the population size that adds up to ignorant users is expressed as γI that are removed from users who consume and adopt disinformation.

We can formulate the SIR compartmental models by introducing more compartments such as hesitants, who can be the users who have not yet decided to adopt the disinformation or not. Although the introduction of new compartments can simulate the real-world information transfer process more accurately, tweaking the parameters of the basic SIR model results in different transformation evolution paths at a relatively lower computational cost. For this reason, the basic SIR model is sufficient to generate different evolutions of users who adopted disinformation at a reasonable computational cost. Estimating the proportion of population who adopted disinformation results in estimating the evolution of commodity demand changes over time horizon of interest in our proposed model.

The evolution of the proportion of these groups over time is modeled on the basis of the homogeneous SIR model, which is mathematically represented by the system of differential equations (a.k.a. mean-field equations) (1)–(3) subject to constraint (4).

$$\frac{dS_{i,t}}{dt} = -\beta S_{i,t} I_{i,t},\ \forall i \in V,\ \forall t \in T, \tag{1}$$

$$\frac{dI_{i,t}}{dt} = \beta S_{i,t} I_{i,t} - \gamma I_{i,t},\ \forall i \in V,\ \forall t \in T, \tag{2}$$

$$\frac{dR_{i,t}}{dt} = \gamma I_{i,t},\ \forall i \in V,\ \forall t \in T, \tag{3}$$

$$S_{i,t} + I_{i,t} + R_{i,t} = 1,\ \forall i \in V,\ \forall t \in T. \tag{4}$$

The index $i \in V$ represents the community surrounding node i, and $t \in T$ denotes time. Under the assumption of homogeneity, users are equally likely to interact with other users. Also, we assumed that no users leave their interactions (e.g., leave social media) during the time of analysis. Therefore, the sum of proportions of the three categories remains constant and equals 1.

The user status can change from one state to another over time. A susceptible (unaware) user encounters an infected (disinformed) user that is infected at a rate β, and a user can move from state I to R by detecting disinformation at a rate γ. That is, in essence, β governs the rate at which disinformation spreads, and γ governs the rate at which disinformed users recover their behavior. Our approach is motivated by interactions on social networks[61]. However, since the parameters of the model, γ and β, govern the rate at which disinformation spreads, other means of social interaction (e.g., TV, radio, web forums) can be taken into account with appropriate rate parameter settings.

There are several ways to solve our system of equations such as the Euler and Runge-Kutta (RK) methods and their derivatives[62]. Each method has advantages and disadvantages in terms of accuracy order and computational cost. The SIR model that we have deployed in our analysis is a non-linear model that needs to be solved numerically by multi-stage algorithms to return the estimates with reasonable accuracy. The forward Euler method is a special case of the RK method, so it solves our problem with relatively low accuracy. At the expense of computational cost, we found the RK algorithm suitable for solving our nonlinear system in terms of accuracy[63,64].

Network flow balance optimization model.

Mathematical programming has proven to be an efficient approach to model and optimize engineered systems and processes[65,66]. Network flow balance optimization can be formulated into a mathematical programming model. There are different ways to formulate network flow optimization problems, however, some formulations are more efficient to solve in terms of complexity[67]. Mathematical programming problems are classified based on the type of decision variables, constraints, and objective functions used in the model. To reduce the computation costs of highly complex problems, there exist some reformulations, which help optimization algorithms to iterate relatively faster or converge to optimal solutions with relatively lower iterations. Among these models, linear programming models are polynomially solvable, while integer and mixed-integer programming models (e.g., the models with integer decision variables) are mostly computationally more expensive to solve[68]. Thus, modeling a problem in linear format is much better in terms of computational complexity. If integer variables need to be included in the model, there are reformulation techniques to convert or divide the models to smaller problems to be solved faster. We formulated the network flow balance optimization problem efficiently and as simple as possible to include a relatively low number of integer decision variables. As a result, we could solve the model iteratively in a reasonable amount of time to compare the results of the optimization problem with respect to different values of model parameters.

We model the infrastructure network as a graph $G(V, E)$, where the set of nodes, V, represents the nodes incorporating demand, supply, and transmission nodes. The set of links, E, represents the links that connect the nodes. There is a link between the nodes if there is a transmission line to transmit the commodity. With the

Notation	Description
Sets	
V	Set of infrastructure network nodes
E	Set of infrastructure network links
T	Set of periods
Parameters	
$\bar{t}_t$	Duration of each period starting from time t to the beginning of its next period
q_{it}	The amount of supply in node $i \in V$ at time $t \in T$
m_{ijt}	Capacity of link from node i to node j at time $t \in T$
p_{it}	Community size surrounding the node i at time $t \in T$
d_{it}^c	Commodity consumption per capita by the community surrounding the node i at time $t \in T$
r_{it}^p	Proportion of commodity consumption of the community surrounding the node $i \in V$ at time $t \in T$ responsive to price shift
ρ_{it}	Estimated sensitivity of commodity consumption of the community surrounding the node $i \in V$ at time $t \in T$ based on the price shift
$\dot{I}_{it}$	Local derivative (change per unit of time interval $\bar{t}_t$) of the proportion of community surrounding node $i \in V$ targeted by disinformation at time $t \in T$
r_{it}	The proportion at which $\dot{I}_{i,t}$ can be changed by spreading counter (good) information for the community surrounding the node $i \in V$ at time $t \in T$
n_t^p	Total number of target locations informed by counter (good) information at time $t \in T$
Decision variables	
x_{ijt}	The amount of transmitted commodity (flow) from node $i \in V$ to node $j \in V$ at time $t \in T$
h_{it}	Shortage (undersupplied) amount of commodity at node $i \in V$ at time $t \in T$
e_{it}	Excess (oversupplied) amount of commodity at node $i \in V$ at time $t \in T$
d_{it}	Nominal demand of commodity in node $i \in V$ at time $t \in T$
I_{it}	Proportion of community surrounding node $i \in V$ at time $t \in T$ adopted disinformation
g_{it}	=1 if counter (good) information is released for the surrounding community in node $i \in V$ at time $t \in T$; =0 otherwise

Table 1. Model notation.

notation found in Table 1, the following is a mixed integer programming (MIP) model to protect the performance of the critical infrastructure network against disinformation dissemination.

$$\min_{x,h,e,d,I,g} \sum_{i \in V,\ t \in T} h_{it} \tag{5}$$

s.t.

$$\sum_{j \in V:(j,i) \in E} x_{jit} - \sum_{k \in V:(i,k) \in E} x_{ikt} + h_{it} - e_{it} + q_{it} - d_{it} = 0, \qquad \forall i \in V, \forall t \in T, \tag{6}$$

$$x_{ijt} \le m_{ijt}, \qquad \forall i \in V, \forall j \in V, \forall t \in T, \tag{7}$$

$$d_{it} = p_{it} d_{it}^c \left\{ (1 - I_{it}) + \{ I_{it} [r_{it}^p (1 + \rho_{it}) + (1 - r_{it}^p)] \} \right\}, \qquad \forall i \in V, \forall t \in T, \tag{8}$$

$$I_{i,t+\bar{t}_t} = I_{it} + \dot{I}_{i,t+\bar{t}_t} (1 - r_{it}\, g_{it}), \qquad \forall i \in V \setminus \{|\ V\ |\}, \forall t \in T, \tag{9}$$

$$\sum_{i \in V} g_{it} \le n_t^p, \qquad \forall t \in T, \tag{10}$$

$$x_{ijt}, h_{it}, e_{it} \in \mathbb{R}_{\ge 0}, \qquad d_{it}, I_{it} \in \mathbb{R}_{\ge 0}, \qquad g_{it} \in \{0, 1\}. \tag{11}$$

Equation (5) is the objective function that minimizes the total amount of commodity shortage resulting from altered consumption behavior over time. Constraint (6) guarantees the balance of the input, output, produced and consumed of the commodity for all nodes. The balance equations are implicitly borrowed from the model proposed by Tang et al.[69]. Constraint (7) limits the capacity of the links. Constraint (8) represents the baseline and responsive demand in terms of the number of users targeted for disinformation given the elasticity of the commodity demand with respect to exogenous factors (e.g., discount price message). Constraint (9) is used to account for the counter- spread of good information as a strategy to control disinformation dissemination.

Constraint (10) limits the number of nodes to focus information countering strategies. The last set of constraints (11) describes the nature of the decision variables.

Solutions to this optimization problem can guide decisions to mitigate an commodity shortage based on disinformation, namely: (i) the amount of commodity flow that should be transmitted through the links, (ii) the optimal shortage or excess in each node, and (iii) the optimal number and location of our communities (surrounding particular nodes) to spread counter information to prevent the adverse effects of disinformation campaigns.

Note that a node cannot experience a shortage and excess simultaneously at the node level. We assume that social media users react to disinformation logically. For example, once a false price discount disinformation is broadcast, social network users consume more commodity relative to their baseline usage.

Case study: performance of the electric power network under disinformation attack

An electric power system is a network of electrical nodes, such as power plants, transformers, or demand points, connected by links that represent transmission lines, cables, or transformers. In such networks, the nodes represent the equipment of the power system and the links are the pathways for the transmission of electrical energies. The electrical energies that are transmitted are called power flows. Electrical power networks are used to satisfy the needs of load nodes (demand nodes) anywhere in the network by transferring the electric power produced by generators (supply nodes) through the links in the network. Each link can carry the maximum commodity through the network, called the flow capacity. Therefore, flow transmissions are limited due to flow capacities in the network. Since the flow capacities are finite in power systems, the problem of transferring the flows to satisfy the demand nodes is a vital network optimization problem that needs to be studied.

To balance the electric power system, several different models and methods are proposed in the literature, such as mathematical optimization and machine learning[70]. For example, Nasrolahpour et al.[71] developed a mixed-integer programming model to alleviate electric power congestion in transmission lines to ultimately minimize the electric power shortage and total cost. Clack et al.[72] developed a linear programming model for electric power balance given its engineering requirements. In the models mentioned above, the common constraint of concern for the authors was a system-wide constraint to guarantee the balance between supply and demand nodes over the network. Also, to obtain realistic solutions to the model, the capacity of the transmission line is specified before optimizing the model. We utilize similar constraints from the literature and include a mechanism to counter disinformation dissemination to defend the spread of disinformation.

In recent years, a handful of papers have begun to address the potential for disinformation to affect commodity consumption. Nguyen et al.[73] developed a vulnerability assessment model to mitigate the adverse effects of disinformation on load shedding. Tang et al.[69] developed an optimization model to minimize total load shedding in a power network under the condition that users react to price disinformation, relating those reactions to user personality traits. Raman et al.[6] developed an attacker-defender optimization model to mitigate strategic urban power distribution system attacks based on price disinformation (e.g., falsely offering prizes for rescheduled power usage) propagated through the community based on the "Believe, Accept, and Follow Through" mechanism.

Among all critical infrastructures, the electric power grid has been attractive to scholars for several reasons: (1) the electric power grid has been at great risk of attack and threats tremendously[74]; (2) electric power grid has been relatively more expensive than other infrastructures[75]; (3) the electric grid is one of the most vital infrastructure during disasters (e.g., Hurricane Sandy in New York) since it is an indirect critical source of commodity for other vital sectors[56]. For these reasons, the analysis of the performance of power grids under dissemination of disinformation has attracted the most attention.

Although we offer a general modeling approach that can be manipulated for a variety of critical infrastructures, our case study is motivated by the electric power grid. Disruptions to power distribution systems can result in substantial economic and social costs[76]. Due to various social factors (e.g., human mistakes, irrationality, intentional gaming, malicious attacks), the electric power grid may become more vulnerable to various kinds of cyber and physical activities when social information becomes tightly integrated into its operation. For example, a coordinated attack could cause a significant impact, such as that experienced in the Ukrainian power grid in 2015[77].

Electric power utilities are increasingly taking advantage of *demand response* programs to reduce or shift electricity usage during peak periods in response to time-based rates or other forms of financial incentives to customers[69,78]. Such demand response programs will be important in the future grid[79,80]. Demand response messaging has been primarily textual, coming from text messages, emails, or other social media messages[6,81]. Naturally, these messages affect human consumption behavior and are used to run an efficient electrical power grid system. Unfortunately, this valuable and effective mechanism could also be used to spread *disinformation*, thus creating a weapon to create a harmful effect - disrupting the power system. Imagine a Tweet being spread by a realistic but fake Twitter account. A discount price is offered to those whose power usage exceeds their average daily use by 30% during summer afternoons. As more and more customers (even those who are not the creators of the disinformation) spread this disinformation and subsequently adopt its message, blackouts will occur more likely due to overloads in the system, along with broader spread impacts to public health and safety. If a threshold of users in a particular geographical location adopts disinformation, a disruption in the power network will occur.

Determining the parameters of the model. The proportion of social media users who may adopt disinformation is found in the SIR model as the proportion of users that make up the group S.

To evaluate the relationship between this spread of disinformation and the demand for electricity, we must also estimate the change in the demand for electricity driven by users whose consumption changed based on

disinformation. The elasticity of the use of electric power measures the responsiveness of the electric power demanded to a change in price. Data from the US state level show that the estimated residential price elasticity of electric power demand is −0.7, suggesting that residential electricity consumption is inelastic to price changes[82,83]. A well-known formula, the midpoint method, used to compute the elasticity of electric power demand is found in Eq. (12), where $pr_{i,t}$ represents the price of the electric power utility at node $i \in V$ at time $t \in T$. As a result, the parameter $\rho_{i,t}$ estimates the proportional change in electric power usage around node $i \in V$ at time $t \in T$ based on a false discount price message.

$$\rho_{i,t} = \frac{d_{i,t+1} - d_{i,t}}{pr_{i,t+1} - pr_{i,t}} \times \frac{pr_{i,t+1} + pr_{i,t}}{d_{i,t+1} + d_{i,t}} \tag{12}$$

Disinformation messages can take different forms to affect electricity use, such as fake weather conditions, fake availability, false prices of alternative sources of electricity (e.g., coal, oil, renewable sources), and false announcements that describe the general economic situation. Such factors have been shown to significantly influence electric power consumption[84,85] with different effects depending on geography and spatiotemporal aggregation[82,83].

The management of loads in electric power systems is highly dependent on the retail price and sales of electricity. Broadcasting false price discount signals on social networks results in increased electricity use by a large number of consumers at once, which may eventually lead to overload or blackout at the node or system level. However, when users who choose to adopt the false pricing message receive it, not all electrical power usage changes accordingly. That is, not all electric power usage is responsive to price changes. For example, price changes might affect a more variable usage appliance (e.g., air conditioner usage may increase) relative to one that is not as variable (e.g., a refrigerator will use the same amount of electricity regardless). Based on the US Energy Information Administration, we assume that 92% of electricity use is responsive to price changes and the rest is consumed to meet basic needs of life[86]. Therefore, we assume that a false discount price message can only affect 92% of electricity usage.

Additionally, not everyone has the ability to receive disinformation messages on social networks. The reports reveal that 82% of the US population were social media users in 2021[87]. Therefore, we assumed that in each population surrounding electric power buses, only 82% people have access to social media directly and are susceptible to receiving the disinformation message. Among these communities, not all residents are sensitive to disinformation about the change in electric price, as not all family members are responsible for household decisions. It is shown that the composition of the home is an influential factor in determining electricity consumption[88]. Thus, we adjust the community p_{it} accordingly to incorporate the effective proportion of the population who may be more responsible for household decisions and who may be more at risk of adopting disinformation.

In summary, we assumed that a fake discount price message can affect 82% of users, and among the total electricity usage they consume, only 92% of their usage may change based on the disinformation they receive (if they adopt the disinformation). As a result, the demand per capita, $d^c_{i,t}$, is modified in the model accordingly.

Designers of electric power transmission systems should ensure that the system can operate normally under unanticipated loads. The safety factor is used to tolerate the system to meet unexpected loads and avoid waste of energy resources. Furthermore, the safety factor of electric power transmission systems is a measure of transmission line reliability that accounts for the possibility of overloaded electric power flow through transmission lines. Due to the lack of data, we use a primary Linear Programming (LP) model with no specified capacity to estimate the capacity of the transmission lines. Assuming that the power distribution lines operate optimally, we take 0.2 (±20%) as a proportion of the allowed volatility of the power flow in the transmission lines. We set the estimated values as the capacity of the transmission lines.

The model we proposed to estimate the capacity of transmission lines is as the following LP problem in Eqs. (13)–(16), where $\bar{d}_{i,t}$ is the nominal demand of the population assigned to the bus i at time t.

$$\min_{x,h,e} \sum_{i \in V,\ t \in T} h_{it} \tag{13}$$

$$\text{s.t.} \tag{14}$$

$$\sum_{j \in V:(j,i) \in E} x_{jit} - \sum_{k \in V:(i,k) \in E} x_{ikt} + h_{it} - e_{it} + q_{it} - \bar{d}_{it} = 0, \qquad \forall i \in V, \forall t \in T \tag{15}$$

$$x_{ijt}, h_{it}, e_{it} \in \mathbb{R}_{\geq 0}. \tag{16}$$

Numerical results. We applied the proposed model to evaluate the effect of disinformation on electric power distribution systems. The electric power nodes supply the electricity to its surrounding community. We overlay the geospatial population data with the topology of the power network to establish the boundary of the model. We use the US Census Application Programming Interface (API) to collect geospatial population data surrounding each electrical power node[89]. The American Community Survey (ACS) provides population data, for use in the model. In addition, we spatially clipped power nodes and links (that is, we generated a shapefile based on spreadsheet data provided by the synthetic power grid data set[90]) that intersected with Los Angeles (LA) County block groups and overlaid it with LA County population data. Approximately 1600 people live in each block group in LA County.

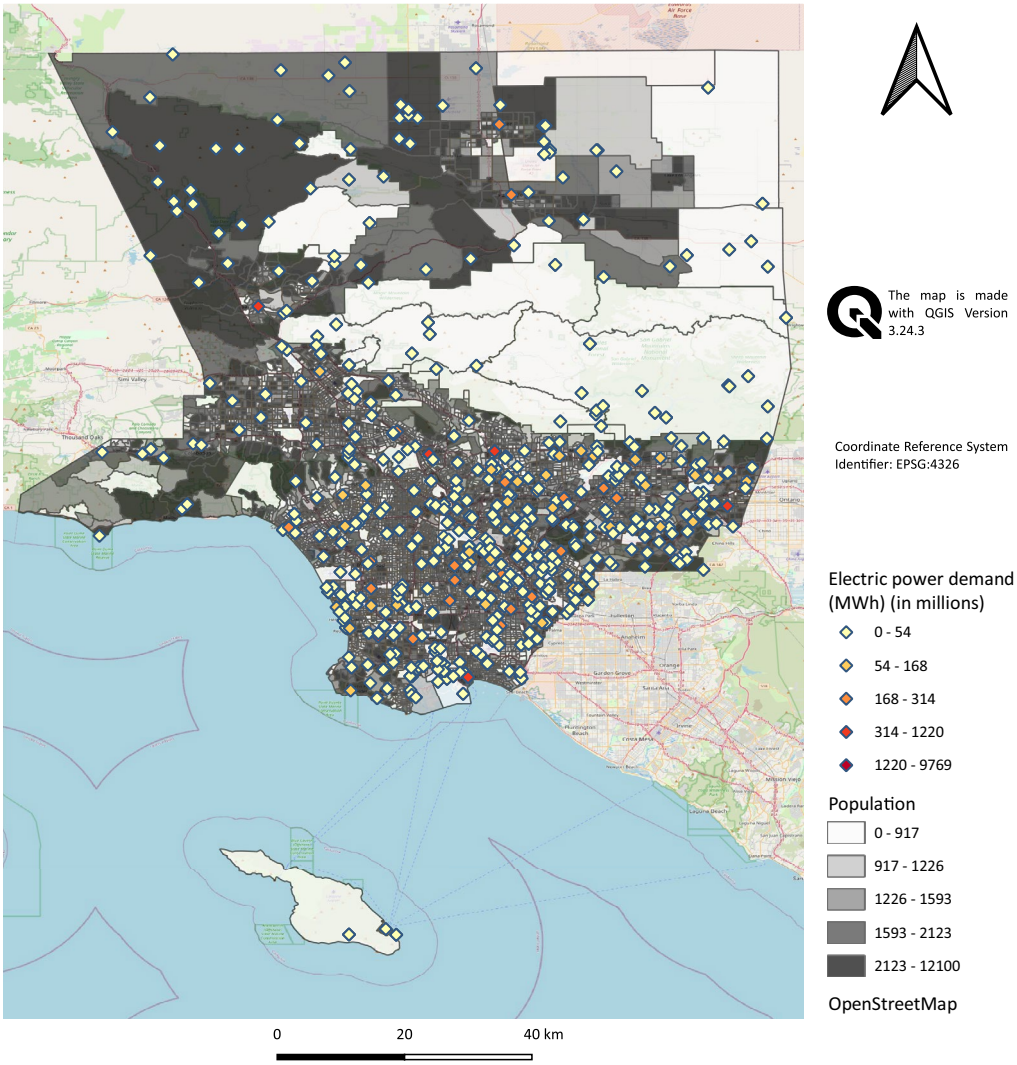


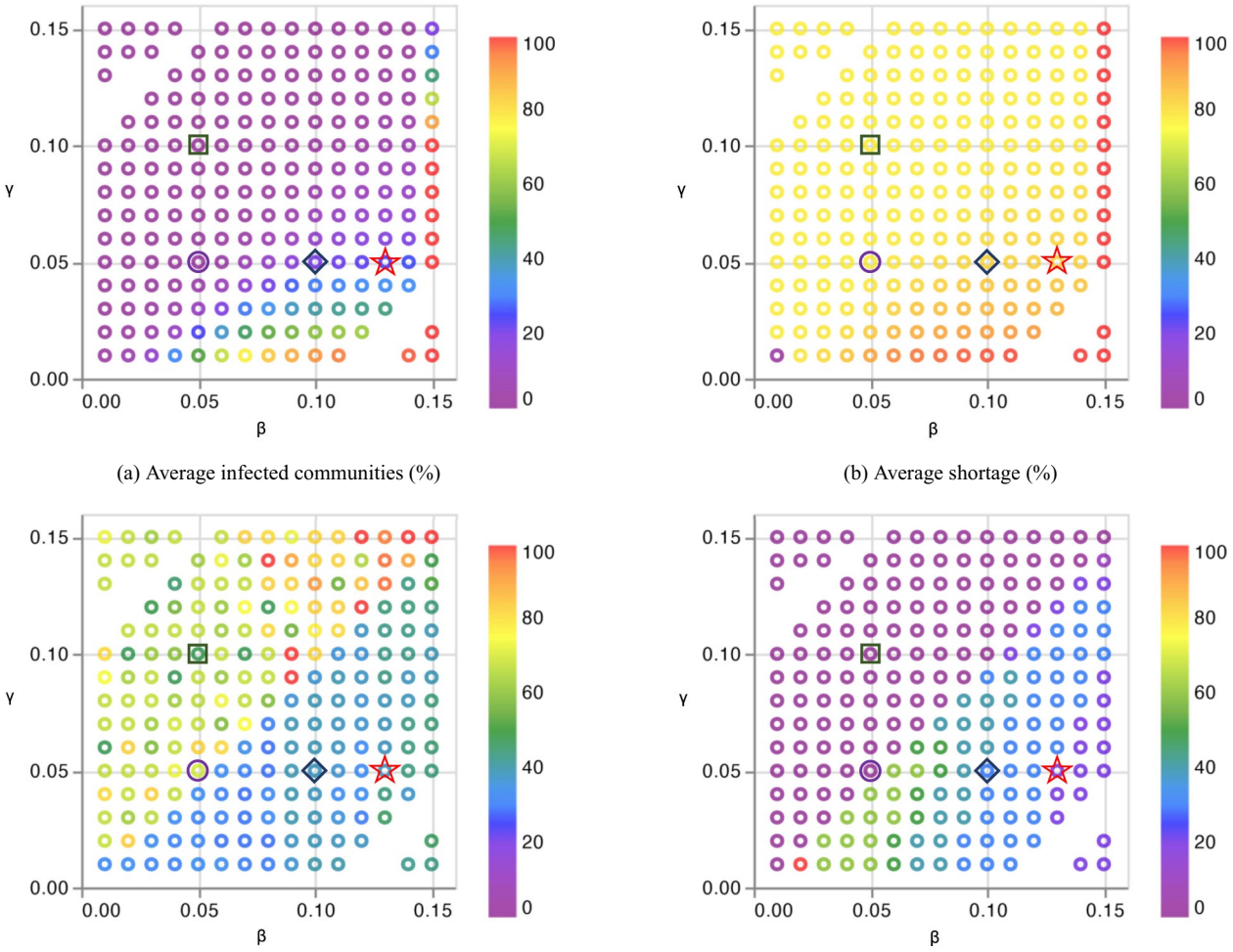
Figure 1. Distribution of electric power demand and population, LA County, USA.

For homeland security purposes, the actual topology of the power grid is not publicly available. However, the information on topology, demand and supply is estimated using *network imitating method based on learning* (NIMBLE)[91], resulting in a publicly accessible synthetic power grid data set for the western interconnection grid of the United States[90]. We limited the boundaries of this larger power distribution system to LA County in California. It incorporates more than 500 power buses (nodes) and 600 transmission lines (links), the geospatial dispersion of which is shown in Fig. 1.

Although studying the effect of disinformation evolution on critical infrastructures at the micro-level makes more sense, highly detailed information is not publicly available. For example, while studying the evolution of disinformation integrated with low-voltage electric power systems may be more desired, we may be more able to estimate parameters of high-voltage electric power systems across a relatively larger geographical area (e.g., the treatment of the impacts of disinformation on a high-voltage network by[92]).

A portion of commodity consumers are assumed to be users of social networks. We relate consumer products to spatially defined block groups, defined as a statistical division of US Census tracts that consists of clusters of blocks that generally contain 600 to 3000 residents of the contiguous area[93], and each group of consumers is linked to the nodes explained.

To relate social media users to electric power buses, we performed geospatial operations for these two features. First, we defined two sets that incorporate electric power buses and aggregated social media users. Electric power buses and social networks are geographically represented by points and polygonal features, respectively. Social media users live in census block groups, defined as a statistical division of US Census tracts that consists of clusters of blocks that generally contain 600 to 3000 residents of the contiguous area[93]. Then we calculated the Euclidean distance matrix between the electrical power buses and the centroid of polygons. As a result, block groups are assigned to one power bus according to their shortest Euclidean distance, and several block groups are mapped to electric power buses and are characterized by estimates of power usage. As such, SIR models are deployed for each social media user within each block group.



(a) Average infected communities (%)

(b) Average shortage (%)

(c) Average flow (%)

(d) Average counter information targets (%)

Figure 2. The results of (**a**) average infected communities, (**b**) average shortage, (**c**) average flows, and (**d**) average counter information targets with different combinations of β and γ as percentage of the baseline value.

Based on the estimated usage of social networks in 2021[87], it is assumed that 82% of the population in each block group have active access to social networks, so they are potentially susceptible to being targeted by disinformation. To run the disinformation propagation model, we considered 1% population being targeted by disinformation at the beginning of the analysis time period. As time goes on, the proportion of susceptible, infected (targeted) and recovered users changes according to the contact rate, the rate of being targeted by disinformation, and the rate of being aware of disinformation.

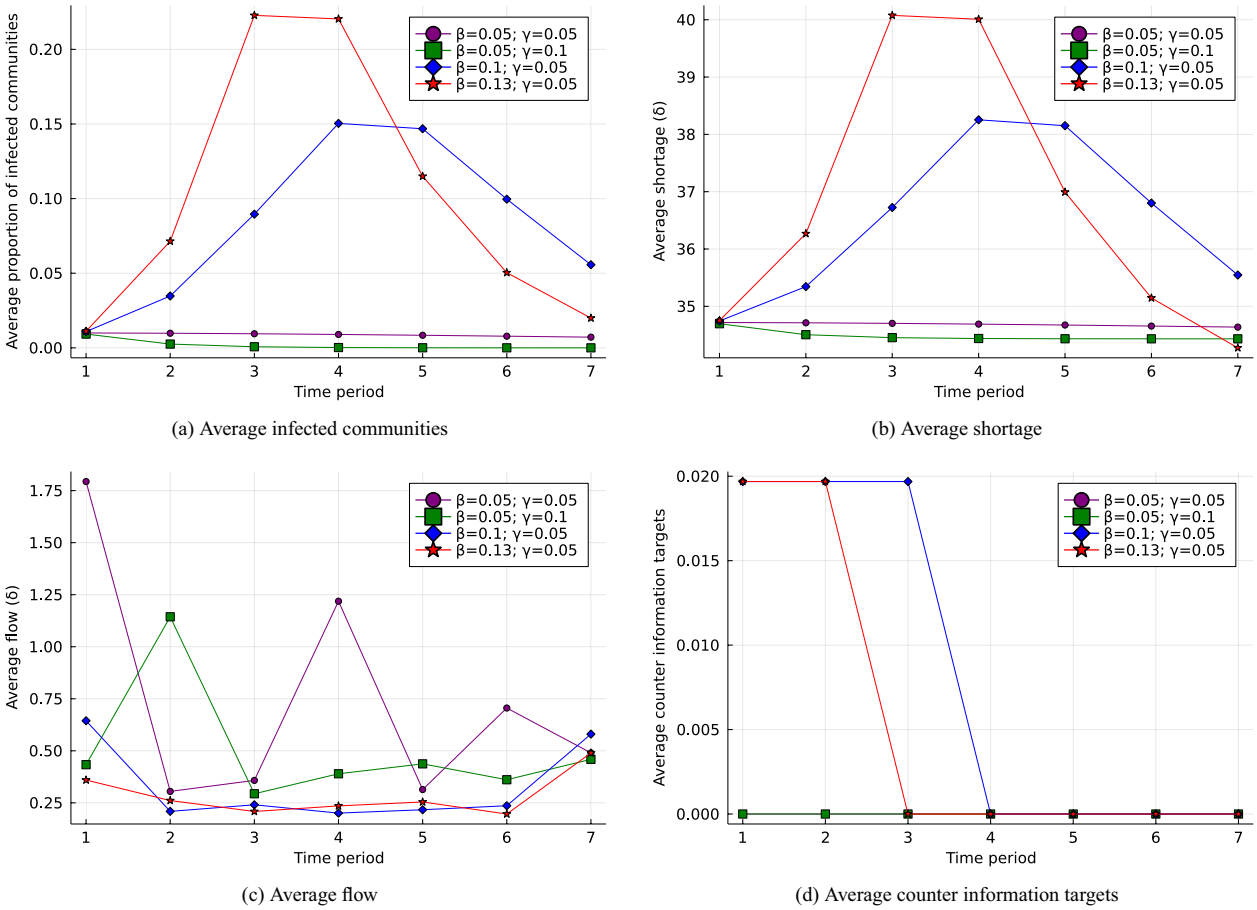
The topology of the power distribution network and the community layer are integrated as a one-to-many setting such that many block groups are assigned to one and only one power bus based on their shortest Euclidean distance to the set of power bus candidates. In other words, the population in block groups is clustered such that the locations of power buses are set as the mean of population clusters.

We interpret the parameters β and γ as the rate of disinformation degree of interest and the rate of awareness, respectively. A higher value of β results in a higher number of people targeted by disinformation per time period. The higher the value of γ, the larger the number of people who become aware of disinformation after being targeted per time period. We analyzed the sensitivity of the solutions for different values of β and γ, as shown in Fig. 2. In these graphs, the ideal condition for β and γ is in the upper left corner of the figures, where the rate of degree of interest takes on the lowest value and the awareness rate is set to its highest value. On the other hand, the worst case is where β is relatively higher and γ is relatively lower, which is located in the lower right corner of the figures. We ran the model with respect to several different instances of the values β and γ to analyze the sensitivity of the total shortage, the total number of communities targeted by counter information, the total flow, and the total infected communities. The results are normalized to show the percentage of difference in the resulting values.

To measure the potential spread of disinformation between social media users, the basic reproduction number ($R_i^0 = \frac{\beta}{\gamma}$) is used. It reveals the expected number of secondary susceptible users that a targeted user can affect with disinformation. For example, given $R_i^0 = 20$, each newly targeted user is expected to affect 20 secondary users within the community i, assuming that all other users contacted are susceptible. To eliminate disinformation or decrease the number of targeted users, the basic reproduction number should satisfy $R_i^0 < 1$, otherwise

Parameters	$\rho_{i,t}$	$r^p_{i,t}$	$r_{i,t}$	n^p_t	$\bar{t}_t$	Horizon
Values	−0.7	0.92	0.2	10	24	169

Table 2. Parameter values.



(a) Average infected communities

(b) Average shortage

(c) Average flow

(d) Average counter information targets

Figure 3. Time series of (**a**) average infected communities, (**b**) average shortage, (**c**) average flow, and (**d**) average counter information targets, with different combinations of β and γ. Note the relationship to particular points in Fig. 2 denoted by shape.

disinformation spreads over the network over time ($R_i^0 > 1$), or the number of targeted users remains constant over time ($R_i^0 = 1$). These concepts help to understand the situations in which the numerical results are analyzed.

We use the differential equation package[94] in Julia programming language to solve the SIR model. We used the mathematical optimization modeling language JuMP[95] to codify our optimization problem in Julia with the optimizer, CPLEX[96], attached to it. Also, we generated the map in Fig. 1 using QGIS Version 3.24.3[97].

Based on the assumptions discussed previously, we run the model several times to evaluate decisions in different situations of spreading disinformation. The values of the parameters we used to run the model are listed in Table 2, and the results with different combinations of β and γ are plotted in Fig. 2. As a validation exercise, we optimize the model for different values of disinformation adoption and detection rate based on Raman et al.[80], who conducted a survey with more than 5000 participants to assess the proportion of people who are expected to adopt and spread disinformation about electricity prices through social networks. They evaluated different scenarios and mapping functions (i.e., linear, quadratic, cubic) to simulate disinformation spread. For simplicity, we adopted the midpoints of the simulated follow-through rates across the mapping functions and performed sensitivity analysis to illustrate the response of the metrics to different scenarios of disinformation propagation in the range. We sampled some notable instances of β and γ combinations (that is, marked by square, circle, diamond, and star shapes) to analyze the evolution of the corresponding metrics over time plotted in Fig. 3.

Figure 2a represents the percentage of infected users in the network with respect to the governing rate of degree of interest in disinformation (β) and awareness (γ). Note that there are some empty spots in this figure (and the rest of the figures) as the SIR model is infeasible for some parameter values. The average number of infected users decreases as the awareness rate increases, although the degree of interest rate is sufficiently high in

most regions. For the lower awareness rate, there is more potential to have infected communities, and it increases further for higher degree of disinformation of interest.

Figure 2b shows the average network shortage with respect to the governing rate of degree of interest and awareness of disinformation. The average shortage increases as the degree of disinformation interest rate increases for a fixed awareness rate. On the other hand, for a higher awareness rate, the average shortage is lower while the degree of disinformation of interest level is fixed. This trend makes intuitive sense, as we saw in Fig. 2a that the average infected communities decrease with higher levels of awareness, and this means that demand increases less caused by disinformation and, therefore, the average shortage is reduced. The effect of the degree of interest in disinformation is greater than the awareness rate, as it always causes a shortage through the network.

Figure 2c represents the average flow in the network with respect to the governing rate of degree of interest and awareness of disinformation. There is a clear limit in the graph where these two rates are equal ($R_i^0 = 1$ or β=γ). On the lower rectangle of values, where $R_i^0 > 1$, the basic reproduction number is high enough to allow disinformation to spread over the network in time, and on the upper rectangle of values, where $R_i^0 < 1$, disinformation has the potential to be eliminated. For the awareness rate above this bound, the average flow is higher in the network, whereas for the region below the bound, we can see much less flow in the network. This is a natural result, as we see that, based on Fig. 2b, the average shortage is lower with greater awareness in the community, and this means that the network can meet demand effectively. In other words, as the average infected communities decreases for a higher awareness rate based on Fig. 2a, the network has more potential to satisfy demands through actual links by transmitting flows.

Figure 2d represents the average number of targeted nodes for counter (good) information with respect to the governing rate of degree of interest and awareness of disinformation. There is a clear limit in the graph where these two rates are equal ($R_i^0 = 1$ or β=γ). For the awareness rate above this bound, the average target counter information is low and not more than 20%. This suggests that if users are at least as aware as the disinformation attracts them, then we can rely less on identifying individuals with whom to supply counter information. This result is in agreement with Fig. 2a and b, as with the higher level of awareness, we have fewer infected communities and also less shortage, which means that less information is needed as a counter mechanism. Furthermore, there are no considerable network shortages that are problematic in this situation. For the lower awareness rate and the higher degree of interest in disinformation, where $R_i^0 > 1$, the model tries to engage more users with counter information, as shown in Fig. 2d.

We compared the evolution of the metrics discussed over time based on a sample combination of values of β and γ. We sampled two instances for $R^0 > 1$, one for $R^0 < 1$, and another for $R^0 = 1$. We use a normalizing constant, δ, as the units of demands and flows of electric power to interpret the output time series plots. With the time series output plots found in Fig. 3a and b, one would expect similar time series for the average infected communities and the average shortage, because an increase in the number of electric power users results in a more substantial power shortage. Figure 3c shows that despite a relatively higher peak of electric power use for $R^0 > 1$, power demands remain unsatisfied as the average flow remains low relative to scenarios $R^0 < 1$ and $R^0 = 1$. Similarly, the average flow over time remains relatively low for $R^0 > 1$, which means that the network capacity has not been used sufficiently to satisfy the demands in the corresponding scenario of disinformation propagation. This interpretation is also clear from Fig. 3a, since the proportion of communities that adopted disinformation is relatively higher than in two other cases. We also observe that flows fluctuate relatively more in scenarios where $R^0 \leq 1$ is average, that is, the use of the network capacity contributed to the satisfaction of demand with a lower spread of disinformation. As Fig. 3d reveals, with a higher intensity of peak demand, we do not necessarily need to target more locations to diffuse counter information. Instead, the duration of disinformation propagation plays a crucial role in selecting the number of communities that become aware of disinformation.

Concluding remarks

The proposed model aims to analyze the adverse effect of disinformation on electric power networks by integrating (i) an epidemiological SIR model to characterize the spread of disinformation in the communities surrounding electric power nodes and (ii) an electric power network optimization model focusing on minimization of the electric power shortage. In particular, we try to mitigate the effects of disinformation by identifying vulnerable power nodes and countering disinformation spread by targeting particular communities with the spread of (good) information. To illustrate the proposed model, we solved a large-scale electric power network problem associated with Los Angeles County, California.

The evaluation of the results of our proposed model reveals how adversaries can interrupt the performance of critical infrastructures to deliver commodities to customers. In addition, we show how the intensity and duration of disinformation diffusion can be monitored to manage infrastructure performance and make communities counter the disinformation. The proposed model opens up a new space for studying the effect of disinformation diffusion in other infrastructures and managing their performance under a disinformation attack. By applying our proposed model to the large-scale electric power network, under several different scenarios of disinformation diffusion throughout Los Angeles County, we showed how our model can be applied to control the propagation of disinformation projected on the performance of infrastructures.

Due to its criticality, the electric power distribution network is used to illustrate the proposed methodology. However, the proposed integration of epidemiological and network flow models is generally applicable to a wide range of infrastructure networks with appropriate changes to the physical infrastructure flow model (e.g., physical laws that restrict the flow in gas pipelines, user behavior that affects traffic flow in a transportation network).

A primary limitation of this model is the boundary we need to draw to select an electric power distribution network to ensure a timely solution to the optimization problem. However, electric power networks are not isolated, as they interact with each other to mitigate shortages in different stations. With the evolution of

computation technology, this model can be applicable and tested on larger-scale networks. Moreover, the proposed model is useful to study the effects of disinformation in other types of critical infrastructure networks, including water and gas, among others, with appropriate physical representations governing the optimization model. Future work includes applying the proposed model to other critical infrastructure networks such as gas distribution systems, nuclear power plants, water distribution systems, etc. To extend the disinformation compartmental model, novel and flexible models (such as agent-based models) can be developed and integrated with the proposed optimization model. Moreover, some parameters used in this article are evaluated by sensitivity analysis or borrowed from previous studies or reports available online. In the future, studies will include a broader range of analysis on the fixed parameters applied to our proposed method. For example, since consumption may not vary linearly in different price ranges during disinformation spread, future work can consider the responsiveness of consumption behavior.

Data availability

The datasets generated and/or analyzed during the current study are available in the Github repository https://github.com/jamalzadeh1400/OU_disinformation/tree/cc6e01365e3159c7d7e7b5b1b65ac1706e37b04f.

Received: 18 February 2022; Accepted: 18 July 2022
Published online: 26 July 2022

References

1. Floridi, L. Is semantic information meaningful data?. *Phil. Phenomenol. Res.* **70**, 351–370 (2005).
2. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
3. Allcott, H., Gentzkow, M. & Yu, C. Trends in the diffusion of misinformation on social media. *Res. Pol.* **6**, 2053168019848554 (2019).
4. Huang, K., Zhou, C., Qin, Y. & Tu, W. A game-theoretic approach to cross-layer security decision-making in industrial cyber-physical systems. *IEEE Trans. Industr. Electron.* **67**, 2371–2379 (2019).
5. Liang, G., He, W., Xu, C., Chen, L. & Zeng, J. Rumor identification in microblogging systems based on users' behavior. *IEEE Trans. Comput. Soc. Syst.* **2**, 99–108 (2015).
6. Raman, G., Peng, J.C.-H. & Rahwan, T. Manipulating residents' behavior to attack the urban power distribution system. *IEEE Trans. Ind. Inf.* **15**, 5575–5587 (2019).
7. Thomaselli, R. Man Tries to Delay Flight by Reporting Fake Bomb Threat. https://www.travelpulse.com/news/airlines/man-tries-to-delay-flight-by-reporting-fake-bomb-threat.html. Accessed: 2020-02-01. (2020).
8. Molina, M. D. & Sundar, S. S. Technological affordances can promote misinformation. *Journal. Truth Age Soc. Med.* 40–57 (2019).
9. Tufnell, N. Students hack Waze, send in army of traffic bots. https://www.wired.co.uk/article/waze-hacked-fake-traffic-jam. Accessed: 2020-02-01. (2014).
10. Barrett, b. An Artist Used 99 Phones to Fake a Google Maps Traffic Jam. https://www.wired.com/story/99-phones-fake-google-maps-traffic-jam/. Accessed: 2020-02-01. (2020).
11. Waniek, M., Raman, G., AlShebli, B., Peng, J.C.-H. & Rahwan, T. Traffic networks are vulnerable to disinformation attacks. *Sci. Rep.* **11**, 1–11 (2021).
12. DeBruhl, B. & Tague, P. Optimizing a misinformation and misbehavior (mib) attack targeting vehicle platoons. 1–5 (2018).
13. Hamill, J. T. Analysis of layered social networks. (Air Force Institute of Technology, 2006).
14. Hamill, J. T., Deckro, R. F., Wiley, V. D. & Renfro, R. S. Gains, losses and thresholds of influence in social networks. *Int. J. Op. Res.* **2**, 357–379 (2007).
15. Meel, P. & Vishwakarma, D. K. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Syst. Appl.* **153**, 112986 (2020).
16. Wardle, C., Derakhshan, H. et al. Thinking about "information disorder": formats of misinformation, disinformation, and mal-information. Ireton, Cherilyn; Posetti, Julie. Journalism,"fake news" & disinformation. Paris: Unesco 43–54 (2018).
17. Santos-Damorim, K. & de Oliveira Miranda, M. K. F. Misinformation, disinformation, and malinformation: Clarifying the definitions and examples in disinfodemic times. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação* **26**, 01–23 (2021).
18. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London Ser. A Contain. Papers Math. Phys. Character* **115**, 700–721 (1927).
19. Sahafizadeh, E. & Ladani, B. T. The impact of group propagation on rumor spreading in mobile social networks. *Phys. A* **506**, 412–423 (2018).
20. Bodaghi, A., Goliaei, S. & Salehi, M. The number of followings as an influential factor in rumor spreading. *Appl. Math. Comput.* **357**, 167–184 (2019).
21. Beskow, D. M. & Carley, K. M. Agent based simulation of bot disinformation maneuvers in twitter. pp 750–761 (2019).
22. Wang, Y., Qing, F., Chai, J. P. & Ni, Y. P. Spreading dynamics of a 2sih2r, rumor spreading model in the homogeneous network. Complexity **2021** (2021).
23. Zhao, L. *et al.* Sihr rumor spreading model in social networks. *Phys. A* **391**, 2444–2453 (2012).
24. Han, Q., Wen, H. & Miao, F. Rumor spreading in interdependent social networks. *Peer-to-Peer Netw. Appl.* **11**, 955–965 (2018).
25. Shrivastava, G. *et al.* Defensive modeling of fake news through online social networks. *IEEE Trans. Comput. Soc. Syst.* **7**, 1159–1167 (2020).
26. Merriam-Webster Dictionary. https://www.merriam-webster.com/dictionary/awareness. Accessed: 2022-01-03.
27. Rui, X., Meng, F., Wang, Z., Yuan, G. & Du, C. Spir: The potential spreaders involved sir model for information diffusion in social networks. *Phys. A* **506**, 254–269 (2018).
28. Hethcote, H. W. The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653 (2000).
29. Keeling, M. J. & Eames, K. T. Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307 (2005).
30. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015).
31. Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I. & Castillo-Chávez, C. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Phys. A* **364**, 513–536 (2006).
32. Woo, J. & Chen, H. Epidemic model for information diffusion in web forums: Experiments in marketing exchange and political dialog. *Springerplus* **5**, 1–19 (2016).
33. Liu, Q., Li, T. & Sun, M. The analysis of an seir rumor propagation model on heterogeneous network. *Phys. A* **469**, 372–380 (2017).
34. Chen, N., Zhu, X. & Chen, Y. Information spreading on complex networks with general group distribution. *Phys. A* **523**, 671–676 (2019).

35. Woo, J., Son, J. & Chen, H. An sir model for violent topic diffusion in social media. 15–19 (2011).
36. Jin, F., Dougherty, E., Saraf, P., Cao, Y. & Ramakrishnan, N. Epidemiological modeling of news and rumors on twitter. 1–9 (2013).
37. Wang, Q., Lin, Z., Jin, Y., Cheng, S. & Yang, T. Esis: Emotion-based spreader-ignorant-stifler model for information diffusion. *Knowl.-Based Syst.* **81**, 46–55 (2015).
38. He, Z., Cai, Z. & Wang, X. Modeling propagation dynamics and developing optimized countermeasures for rumor spreading in online social networks. 205–214 (2015).
39. Khurana, P. & Kumar, D. Sir model for fake news spreading through whatsapp. 26–27 (2018).
40. Zhao, L., Wang, J. & Huang, R. Immunization against the spread of rumors in homogenous networks. *PLoS ONE* **10**, e0124978 (2015).
41. Zhao, Z. *et al.* Fake news propagates differently from real news even at early stages of spreading. *EPJ Data Sci.* **9**, 1–14 (2020).
42. Barker, K. *et al.* Defining resilience analytics for interdependent cyber-physical-social networks. *Sustain. Resil. Infrastruct.* **2**, 59–67 (2017).
43. Hosseini, S., Barker, K. & Ramirez-Marquez, J. E. A review of definitions and measures of system resilience. *Reliab. Eng. Syst. Saf.* **145**, 47–61 (2016).
44. Liu, W. & Song, Z. Review of studies on the resilience of urban critical infrastructure networks. *Reliab. Eng. Syst. Saf.* **193**, 106617 (2020).
45. Cheng, J., Liu, Q., Hui, Q. & Choobineh, F. The joint optimization of critical interdependent infrastructure of an electricity-water-gas system. pp 61–73 (2019).
46. Hsu, N.-S. & Cheng, K.-W. Network flow optimization model for basin-scale water supply planning. *J. Water Resour. Plan. Manag.* **128**, 102–112 (2002).
47. Tahiri, A., Ladeveze, D., Chiron, P., Archimede, B. & Lhuissier, L. Reservoir management using a network flow optimization model considering quadratic convex cost functions on arcs. *Water Resour. Manage* **32**, 3505–3518 (2018).
48. Martin, A., Möller, M. & Moritz, S. Mixed integer models for the stationary case of gas network optimization. *Math. Program.* **105**, 563–582 (2006).
49. Banda, M. K., Herty, M. & Klar, A. Gas flow in pipeline networks. *Netw. Heterog. Media* **1**, 41 (2006).
50. Csikós, A., Charalambous, T., Farhadi, H., Kulcsár, B. & Wymeersch, H. Network traffic flow optimization under performance constraints. *Transp. Res. Part C: Emerg. Technol.* **83**, 120–133 (2017).
51. Darayi, M., Barker, K. & Santos, J. R. Component importance measures for multi-industry vulnerability of a freight transportation network. *Netw. Spat. Econ.* **17**, 1111–1136 (2017).
52. Vasin, A., Grigoryeva, O. & Tsyganov, N. A model for optimization of transport infrastructure for some homogeneous goods markets. *J. Global Optim.* **76**, 499–518 (2020).
53. Costa, A., Georgiadis, D., Ng, T. S. & Sim, M. An optimization model for power grid fortification to maximize attack immunity. *Int. J. Electr. Power Energy Syst.* **99**, 594–602 (2018).
54. Leuthold, F. U., Weigt, H. & von Hirschhausen, C. A large-scale spatial optimization model of the European electricity market. *Netw. Spat. Econ.* **12**, 75–107 (2012).
55. Wirtz, M., Hahn, M., Schreiber, T. & Müller, D. Design optimization of multi-energy systems using mixed-integer linear programming: Which model complexity and level of detail is sufficient?. *Energy Convers. Manage.* **240**, 114249 (2021).
56. Haraguchi, M. & Kim, S. Critical infrastructure interdependence in New York city during hurricane sandy. *Int. J. Disaster Resil. Built Environ.* (2016).
57. González, A. D., Dueñas-Osorio, L., Sánchez-Silva, M. & Medaglia, A. L. The interdependent network design problem for optimal infrastructure system restoration. *Comput. Aided Civ. Infrastruct. Eng.* **31**, 334–350 (2016).
58. Almoghathawi, Y., González, A. D. & Barker, K. Exploring recovery strategies for optimal interdependent infrastructure network resilience. *Netw. Spat. Econ.* **21**, 229–260 (2021).
59. Ghorbani-Renani, N., González, A. D., Barker, K. & Morshedlou, N. Protection-interdiction-restoration: Tri-level optimization for enhancing interdependent network resilience. *Reliab. Eng. Syst. Saf.* **199**, 106907 (2020).
60. Ouyang, M. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliab. Eng. Syst. Saf.* **121**, 43–60 (2014).
61. Watts, D. J., Rothschild, D. M. & Mobius, M. Measuring the news and its impact on democracy. *Proc. Natl. Acad. Sci.* **118**, e1912443118 (2021).
62. Dormand, J. R. & Prince, P. J. A family of embedded runge-kutta formulae. *J. Comput. Appl. Math.* **6**, 19–26 (1980).
63. Medvedeva, M., Simos, T. E., Tsitouras, C. & Katsikis, V. Direct estimation of sir model parameters through second-order finite differences. *Math. Methods Appl. Sci.* **44**, 3819–3826 (2021).
64. Tsitouras, C. Runge-kutta pairs of order 5 (4) satisfying only the first column simplifying assumption. *Comput. Math. Appl.* **62**, 770–775 (2011).
65. Luenberger, D. G. *Optimization by Vector Space Methods* (John Wiley & Sons, 1997).
66. Bertsekas, D. *Convex Optimization Algorithms* (Athena Scientific, Berlin, 2015).
67. Bertsekas, D. *Network Optimization: Continuous and Discrete Models* (Athena Scientific, 1998).
68. Li, W. *et al.* Parameterized algorithms of fundamental np-hard problems: A survey. *HCIS* **10**, 1–24 (2020).
69. Tang, D., Fang, Y.-P., Zio, E. & Ramirez-Marquez, J. E. Resilience of smart power grids to false pricing attacks in the social network. *IEEE Access* **7**, 80491–80505 (2019).
70. Fang, X., Misra, S., Xue, G. & Yang, D. Smart grid'the new and improved power grid: A survey. *IEEE Commun. Surv. Tutor.* **14**, 944–980 (2011).
71. Nasrolahpour, E., Ghasemi, H. & Khanabadi, M. Optimal transmission congestion management by means of substation reconfiguration. pp 416–421 (2012).
72. Clack, C., Xie, Y. & MacDonald, A. Linear programming techniques for developing an optimal electrical system including high-voltage direct-current transmission and storage. *Int. J. Electr. Power Energy Syst.* **68**, 103–114 (2015).
73. Nguyen, L. N., Smith, J. D. & Thai, M. T. Vulnerability assessment of social-smart grids: An algorithmic approach. pp 1–7 (2019).
74. Tian, W. *et al.* Prospect theoretic study of honeypot defense against advanced persistent threats in power grid. *IEEE Access* **8**, 64075–64085 (2020).
75. Kovendan, A. & Sridharan, D. Development of smart grid system in India: A survey. pp 275–285 (2017).
76. Garcia Tapia, A., Suarez, M., Ramirez-Marquez, J. . E. & Barker, K. Evaluating and visualizing the economic impact of commercial districts due to an electric power network disruption. *Risk Anal.* **39**, 2032–2053 (2019).
77. Zetter, K. Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid. Wired. https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/. Accessed: 2021-12-10. (2016).
78. Lund, P. D., Lindgren, J., Mikkola, J. & Salpakari, J. Review of energy system flexibility measures to enable high levels of variable renewable electricity. *Renew. Sustain. Energy Rev.* **45**, 785–807 (2015).
79. Schuitema, G., Ryan, L. & Aravena, C. The consumer's role in flexible energy systems: An interdisciplinary approach to changing consumers' behavior. *IEEE Power Energ. Mag.* **15**, 53–60 (2017).
80. Raman, G., AlShebli, B., Waniek, M., Rahwan, T. & Peng, J.C.-H. How weaponizing disinformation can bring down a city's power grid. *PLoS ONE* **15**, e0236517 (2020).
81. Jain, M. et al. Methodologies for effective demand response messaging. pp 453–458 (2015).

82. Miller, M. & Alberini, A. Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from us data. *Energy Policy* **97**, 235–249 (2016).
83. Burke, P. J. & Abayasekara, A. The price elasticity of electricity demand in the united states: A three-dimensional analysis. *Energy J.* **39** (2018).
84. Borenstein, S. To what electricity price do consumers respond? residential demand elasticity under increasing-block pricing. *Prelim. Draft April* **30**, 95 (2009).
85. Wang, B. *et al.* Electricity price and habits: Which would affect household electricity consumption?. *Energy Build.* **240**, 110888 (2021).
86. EIA Website. https://www.eia.gov/energyexplained/use-of-energy/homes.php. Accessed: 2021-11-11.
87. Statista Website. https://www.statista.com/statistics/273476/percentage-of-us-population-with-a-social-network-profile/. Accessed: 2021-12-05.
88. Brounen, D., Kok, N. & Quigley, J. M. Residential energy use and conservation: Economics and demographics. *Eur. Econ. Rev.* **56**, 931–945 (2012).
89. US Census Application Programming Interface. https://www.census.gov/data/developers/data-sets.html. Accessed: 2021-12-11.
90. Wireless and Mobile Networking Lab. https://wimnet.ee.columbia.edu/portfolio/synthetic-power-grids-data-sets/. Accessed: 2021-12-07.
91. Soltan, S., Loh, A. & Zussman, G. A learning-based method for generating synthetic power grids. *IEEE Syst. J.* **13**, 625–634 (2019).
92. Tang, D., Fang, Y. P., Zio, E. & Ramirez-Marquez, J. E. Resilience of smart power grids to false pricing attacks in the social network. *IEEE Access* **7**, 80491–80505 (2019).
93. United States Census Bureau. https://www.census.gov/. Accessed: 2021-12-13.
94. DifferentialEquations.jl: Scientific Machine Learning (SciML) Enabled Simulation and Estimation. https://diffeq.sciml.ai/stable/. Accessed: 2022-02-15.
95. Dunning, I., Huchette, J. & Lubin, M. Jump: A modeling language for mathematical optimization. *SIAM Rev.* **59**, 295–320 (2017).
96. IBM ILOG CPLEX Optimizer. https://www.ibm.com/analytics/cplex-optimizer. Accessed: 2022-02-15.
97. QGIS. https://www.qgis.org/en/site/. Accessed: 2022-06-01.

Acknowledgements

This research was partially funded by the National Institute of Standards and Technology (NIST) Center of Excellence for Risk-Based Community Resilience Planning through a cooperative agreement with Colorado State University [70NANB20H008 and 70NANB15H044]. Also, this research was partially funded by National Science Foundation (NSF) through award 2052930. The contents expressed in this paper are the views of the authors and do not necessarily represent the opinions or views of NIST or the NSF.

Author contributions

S.J. conceived the proposed model and conducted the experiments. K.B., A.D.G., and S.R. analyzed the results and supervised the research. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.