Available online at www.sciencedirect.com
ScienceDirect
Energy Reports 00 (2024) 1–15
Energy Reports

Modeling and Monitoring Social Media Dynamics to predict Electricity Demand Peaks

Isabella Nunes Grieser[c], Tobias Gebhard[a,*], Andrea Tundis[a], Jens Kersten[b], Tobias Elßner[b], Florian Steinke[c]

[a]*Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Mornewegstraße 30, 64293 Darmstadt, Germany*
[b]*Institute of Data Science, German Aerospace Center (DLR), Mälzerstraße 3-5, 07745 Jena, Germany*
[c]*Energy Information Networks and Systems Lab, Technical University of Darmstadt, Landgraf-Georg-Str. 4, 64283 Darmstadt, Germany*

Abstract

Information spread on social media can lead to sudden, synchronized actions. If this affects electricity demands, it could result in critical consequences for the power grid. With the rise of social media and fake news and the increasing adoption of power-intensive devices, the risk of misinformation attacks by manipulating consumer behavior becomes more relevant. This paper presents a novel approach for modeling the potential impact of social media dynamics on power systems. We present a conceptual monitoring framework for the real-time detection of critical information propagation and the short-term prediction of electricity demand peaks. Based on a social network graph, a stochastic epidemiological model, the Susceptible-Infectious-Recovered (SIR) model, is employed to simulate the "viral" spread of information. To estimate model parameters from real data, an optimization algorithm is developed. Twitter data of a past disaster event is acquired and used to create a generalized propagation dynamics model, which can then be used to analyze the impact of altered power demands. Specifically, we simulate a demand response attack, where households receive misinformation about reduced electricity prices, encouraging them to activate appliances. The results demonstrate that the synchronized behavior of a minority of affected consumers can lead to sudden increases in the aggregated demand, significantly surpassing usual demand levels. Furthermore, we examine the peak demand for electric vehicle (EV) charging at different adoption rates, showing that the consequences of synchronized behavior are amplified. Our innovative approach opens up new possibilities for power grid nowcasting and enhancing critical infrastructure resilience in a proactive manner, which can avoid load shedding.

Keywords: Power Grid Monitoring, Power System Resilience, Misinformation Attack, Demand Response, SIR Model, Social Media Data

1. Introduction

Power systems are vulnerable to high-impact, low-probability (HILP) events, which can cause severe disruptions with catastrophic societal consequences [1, 2]. In rare cases, collective human actions can result in synchronized power demands, potentially leading to extreme load spikes in a short time [3, 4]. The COVID-19 pandemic highlighted how unexpected synchronized consumer behavior, such as panic buying, can disrupt critical infrastructures, such as supply chains [5, 6]. Social media can play a major role in such phenomenons, as information can disseminate rapidly and influence users to alter their behavior [5, 7]. With the rise of misinformation and fake news in the recent years, new scenarios and challenges emerge [8]. If a substantial number of people receives and forwards (mis-)information that

*Corresponding author.
E-mail address: tobias.gebhard@dlr.de
Received 27 August 2024.

somehow relates to a change in power demand, the synchronized loads could exceed critical limits and the sudden actions could endanger the stability of the power system, especially if the change is not expected. As resilience means to consider possibly unprecedented HILP events, for future resilient power grids it is essential to understand and anticipate such emerging scenarios and analyze potential consequences for supporting policymakers and implementing countermeasures [9].

By manipulating consumer behavior with targeted misinformation, i.e. disinformation, social attacks on power systems pose a new type of threat. For instance, demand response (DR) systems can create detrimental effects when attacked. Malicious actors could spread misinformation about reduced electricity prices, which could encourage a significant number of consumers to shift energy-intensive activities to a certain time window, when the grid is actually stressed [10, 11]. According to a survey, a notable portion of the population would act on and forward a fake message about an electricity rate discount [12], underscoring the danger of such attacks. With the increasing adoption of participation in DR programs, the propensity to believe such information might increase [11, 12]. Beside such misinformation attacks, the spread of true information could affect critical infrastructures as well, especially during crisis situations [13, 14]. For example, in the case of an evacuation due to a disaster, such as a wildfire, residents might charge their electric vehicles (EVs) immediately [2]. With the increasing adoption of EVs, the simultaneous charging could create excessive power loads [2, 15].

Such "infection-like" spreads in user actions present new challenges for grid operators, who must be aware of spontaneous demand changes to avoid localized or widespread outages. For ensuring grid stability, reliable short-term load predictions are essential. Traditional forecasting approaches typically rely on autoregressive methods that incorporate features with regular patterns, such as historical load, weather conditions, and calendar information [16, 17]. Most of these models operate on time resolutions of an hour or more. Intra-hourly trends, i.e. "very short term" forecasting, are rarely considered [16]. To anticipate the impact of anomalous events, such as spontaneous collective user behavior, new approaches are necessary.

There exist significant correlations between power demand and social media data [18, 19, 20] or search engine trends [21]. It has been shown that large-scale textual and geospatial social media data can approximate electricity utilization patterns well and even beat established forecasting models [18]. Using social media to anticipate changes in demand patterns presents a new opportunity without the need for widespread physical hardware systems such as smart meters. Moreover, the monitoring of activities beyond electrical systems enables the ability to detect critical trends even before they become visible in the electrical domain. Epidemiological models can provide a novel perspective by representing the "viral" propagation of misinformation in social networks [22, 23, 24]. Just as these models trace infection patterns, they can be adapted to simulate how information and behavior spread could influence power consumption.

In this paper, we present a novel approach to modeling and monitoring the impact of social media dynamics on power systems. To represent the spread of misinformation, an epidemiological Susceptible-Infectious-Recovered (SIR) model is employed. In contrast to [25], we use the SIR model in a graph-based and stochastic form. The propagation model is connected to an underlying social network, for which a scale-free graph is used to approximate real-world social connections. To demonstrate the impact of misinformation on power systems, we apply the model in a case study using the scenario of a false pricing attack on DR. The results show that even with a small share of participants, a massive and sudden increase in power demand occurs, which could surpass grid capacity limits. Additionally, we examine the peak demand if EV charging is conducted at different adoption rates, showing to severely amplify the impact of synchronized behavior. Even if the alteration of power demand is limited to household devices in residential areas, the synchronized use and the unexpected action do turn out to be critical.

Unlike previous studies, we leverage real social media data to calibrate our information propagation model. Since data on events related to misinformation propagation, especially with a connection to power systems, are rare, we use a novel data-driven approach. By assuming universal patterns in the form of information propagation on social media, data on a past disaster event is utilized to refine the propagation model. The dataset used for this work contains keyword-filtered tweets related to a forest fire near Berlin [26]. An optimization algorithm is developed for the estimation of model parameters from the number of social media posts over time to generate similar infection progressions with the model. Thereby, we connect the graph-based SIR model with the system-level SIR model based on differential equations. For the given dataset, the infection progress can be well reconstructed with the obtained parameters, except for a drop during night. The calibrated propagation model can represent misinformation events in a general way and be used for the analysis of potential scenarios, e.g. impact assessment of future technologies, as the

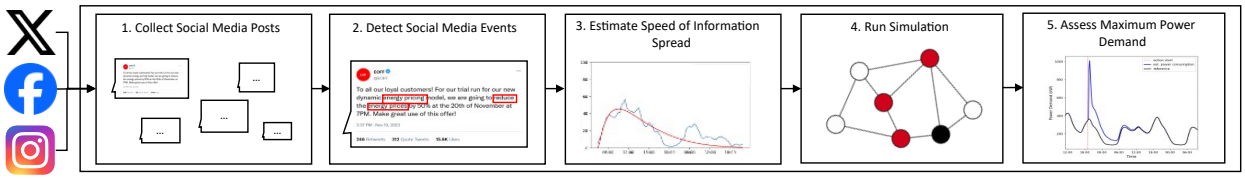


Figure 1: Overview of the suggested monitoring framework

propagation model is decoupled from the power demand model. This can contribute to power system resilience by enabling policy makers to prepare for various scenarios and formulate respective countermeasures.

1.1. Proposal for a Monitoring Framework

The ability to detect critical demand trends in advance is vital for mitigating the consequences and maintaining grid stability. In the event of excessive power consumption, usually *reactive* measures such as load shedding are considered to reduce the strain on the power system. However, load shedding is problematic due to its severe consequences [27] and should be only the last resort to prevent the power grid from collapsing. Current research mostly focused on how to minimize load shedding, but did not explore the detection of critical misinformation for power systems in advance.

With this work, we want to come to a paradigm shift by considering *proactive* methods. The timely detection of critical consumer behavior would increase the time window for grid operators to take effective countermeasures before an event affects the power grid in a critical way and therefore could mitigate the impact and prevent load shedding and blackouts. Moreover, the early detection of an attack would dissuade attackers [11]. To the best of our knowledge, this work is the first combining misinformation propagation modeling and real-time monitoring with social media data to anticipate critical power loads in one framework, which is presented in the following.

The methodology developed in this paper can be integrated in a higher-level conceptual framework for monitoring social media usage in real-time, detecting misinformation propagation events that are critical for power systems, and predicting the short-term peak demand. The proposed framework is comprised of five main steps, with the basic structure illustrated in Figure 1.

1. Social media posts are collected, evaluated, and filtered in real-time. If the impact for a specific geographic region shall be assessed, the location of users may be relevant and needs to be extracted or estimated. Similarly, other online data, e.g. search engine trends, can be acquired.
2. The collected data is analyzed to detect any information that could relate to a change in power demand. This can be done for example by analyzing the semantic information of collected posts using keyword searches or natural language processing. Potential scenarios related to power grids could be defined in advance and be detected given the observations, e.g. by using Bayesian networks [28].
3. The SIR model needs to be initialized with parameters, for which we consider two options. In the *online estimation*, the speed and progress of the ongoing information propagation can be inferred by estimating model parameters with the collected social media data in real-time. The *offline estimation* refers to the use of universal model parameters, fitted with social media data from past events related to information propagation.
4. Based on the identified scenario and the obtained parameters, a simulation of the information propagation and power demand model is executed. As the model contains stochastic elements, multiple simulation runs can provide a comprehensive assessment, considering uncertainty.
5. The simulation results are evaluated to determine the resulting time and amplitude of the aggregated load peak. This "very short-term" prediction can provide stakeholders with useful information to anticipate irregular and critical load deviations on the power system.

This novel monitoring concept could be utilized by grid operators or other relevant authorities to better anticipate and react to critical power demand shifts in a timely manner, mitigating the impact of such events. The active monitoring of social networks and online trends can assist power grid operations in supplementing traditional load forecasting and improving situational awareness, giving grid operators the chance for effective countermeasures. For example, the generation reserve can be activated early enough or precise smart meter curtailments can be planned. This can avoid

load shedding and reduce economic costs associated with such events. Moreover, when critical information is detected, the operators of social networks could be warned to slow down or prevent the dissemination before reaching critical levels.

This paper focuses on the modeling of information propagation and power demand (step 4 and 5) and the offline estimation of model parameters (step 3). We do not cover the collection of social media posts and scenario detection in real-time.

The remainder of the paper is structured as follows. In Section 2, the background and related work are outlined. Section 3 presents our model for information propagation on social media and its impact on power demand. In Section 4, an optimization algorithm is derived to estimate model parameters with data. A case study is presented in Section 5, where the proposed model is simulated and the effects on the power system are analyzed. Additionally, the parameter estimation is applied using a dataset of social media posts. Section 6 provides a discussion on the approach and a comparison with other works. Conclusions are drawn in Section 7.

2. Related Work

In this section, we provide an overview of the theoretical background for this work and related research that connects social media and power systems.

Social networks are often modeled using random graphs. Online social networks tend to exhibit scale-free characteristics, small-world properties, and a high clustering coefficient [29]. A variety of random graph models have been developed with the objective of reproducing these characteristics. Notable examples include the Watts–Strogatz graph [30] and the Barabási–Albert graph [31]. Both were also used in the analysis of the spread of information [22, 12, 24].

The spread of information and misinformation in social networks can be modeled with epidemiological models [22], where the "viral" propagation of information can be seen as an infection that spreads through a social network. A commonly used epidemiological model is the Susceptible-Infectious-Recovered (SIR) model. In its compartmental set-up, entities are assigned to three different states, representing the stages of the infection process [22]. The behavior of each entity is contingent upon its state. In the context of modeling information propagation processes, the states refer to the following meaning:

- **Susceptible (S)** entities are unaware of the information being spread and are receptive to the information.
- **Infected (I)** entities know and act on the information, sharing it with their acquaintances.
- **Recovered (R)** entities know of the information, but do not act on it.

There exist several extensions of the SIR model that refine the states which entities can assume during the infection process. Notable examples include the Susceptible-Infected-Recovered-Susceptible (SIRS) model, the Susceptible-Exposed-Infected-Recovered (SEIR) model [22], and the Susceptible-Exposed-Infected-Skeptic (SEIZ) model [23].

Most SIR models are based on differential equations $\frac{dS}{dt}, \frac{dI}{dt}, \frac{dR}{dt}$, and represent the total number of entities in each state S, I, R in the system at a given time t. These models provide a top-down (system-level) view of the infection process and do not model entities individually. In contrast, graph-based SIR models consider a graph, representing the possible connections where information can flow, and compute the state of each individual entity given their neighbors' states, constituting a bottom-up approach. These models can incorporate stochastic state equations, which often yields more realistic results compared to deterministic models [32].

In the field of (mis-)information propagation, the potential impact on power systems has recently been in focus, most commonly considering false pricing attacks [33, 34, 10, 11, 12, 25]. The information propagation models employed in the studies cover independent cascade models [33, 34, 12], threshold-based influence models [10, 12], and the SIR model in its system-level form [25, 14]. Graph-based SIR models have rarely been applied in the context of critical infrastructures yet, for example to model the spreading of computer viruses [35].

As the change in power demand induced by the social media (mis-)information highly depends on the concrete type of situation, the power demand model needs to consider the specific scenario. For the false pricing attack, an increase in demand by a constant factor [33] or a dedicated additional power demand on top to a regular demand [10] was assumed.

However, none of the aforementioned studies calibrated or validated their models by using real social media data. This is due to the fact that data on such events is rarely available. In this work, we aim for using social media data from similar propagation events to enhance information propagation models.

As a countermeasure for such situations, previous work mostly considered reactive measures, such as load shedding [33, 10, 34, 25]. The minimization of load shedding can be formulated as a mixed integer program (MIP) if the power system is represented as a balanced network flow model [25, 10]. Preventive methods, e.g. real-time monitoring to predict demand peaks before their occurrence, have not been considered.

A game-theoretic analysis of a strategic demand response attack against an electric utility operator as defender has shown that the time of detection plays a significant role in dissuading the attacker to reduce the impact or prevent the attack [11], underscoring the need for the timely detection of such events.

3. Information Propagation and Power Demand Model

This section presents our model of information propagation in social networks and its impacts on power demand. The simulation of the model corresponds to part 4 of the monitoring framework (see Figure 1).

The model consists of three main components. First, we introduce the social network model representing the social connections between the entities, e.g. individuals, as a graph. Second, we discuss the modeling of the information propagation through the social network by using a modified version of the SIR model in a graph-based and stochastic form. Third, the modeling of the resulting changes in the power demand of affected entities is described. A summary of all model parameters is given in Table 1.

3.1. Social Network

We model the social network as a graph $G = (V, E)$, where entities are defined as nodes V and social connections between them are represented as edges E. Since the real social structures of a population are difficult to obtain, we use random graphs with characteristics similar to real social networks.

In this work, we use the Barabási–Albert (BA) graph [31]. The graph $\mathrm{BA}_{N,k}$ uses two parameters. N is the number of nodes in the graph and $k \in [1, N)$ is the number of edges which are generated when a new node is added to the graph. Thus, k is an indicator for the connectivity of the social network.

3.2. Information Propagation

The social network graph resembles the possible routes where information can propagate. If the edge $e_{i,j}$ exists, information can spread from an entity i to entity j with a probability p which is computed with the SIR model.

A graph-based SIR model, similar to [24] is used in this work to model the propagation of information. To represent uncertainty in the stochastic process of information propagation, the model uses probabilistic state transitions. The model differs from SIR models used for epidemiological analysis by the possibility to transition from *Susceptible* directly to *Recovered*, representing people that are skeptical to the information right away. The state diagram of the model is shown in Figure 2.

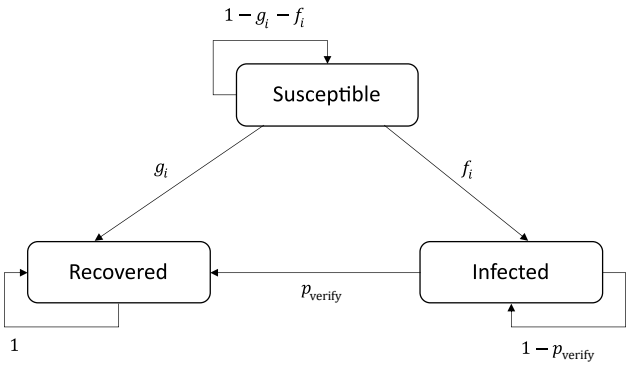


Figure 2: State chart of the modified information propagation model according to (1)

The state of an entity i at the time step t_n is defined as $s_i(t_n) = \left[s_i^S(t_n), s_i^I(t_n), s_i^R(t_n)\right]$, where $s_i^X(t_n) = 1$ if i is in state X, otherwise 0. An entity i may transition to another state X at time step t_{n+1} with the probability $p_i^X(t_{n+1})$ as

$$p_i^S(t_{n+1}) = (1 - f_i(t_n) - g_i(t_n))\, s_i^S(t_n) \tag{1a}$$

$$p_i^I(t_{n+1}) = f_i(t_n) s_i^S(t_n) + (1 - p_{\text{verify}}) s_i^I(t_n) \tag{1b}$$

$$p_i^R(t_{n+1}) = g_i(t_n) s_i^S(t_n) + p_{\text{verify}}\, s_i^I(t) + s_i^R(t_n) \tag{1c}$$

with

$$f_i(t_n) = \begin{cases} 0 & \text{if } h_i^I(t_n) = h_i^R(t_n) = 0 \\ \beta \frac{(1+\alpha) h_i^I(t_n)}{(1+\alpha) h_i^I(t_n) + (1-\alpha) h_i^R(t_n)} & \text{else} \end{cases} \tag{2a}$$

$$g_i(t_n) = \begin{cases} 0 & \text{if } h_i^I(t_n) = h_i^R(t_n) = 0 \\ \beta \frac{(1-\alpha) h_i^R(t_n)}{(1+\alpha) h_i^I(t_n) + (1-\alpha) h_i^R(t_n)} & \text{else} \end{cases} . \tag{2b}$$

At the beginning of the infection process ($n = 0$), all entities are in the *susceptible* state, except for one or multiple *infected* entities, which are considered the source of the misinformation. A *susceptible* entity i may change its state to either *infected* or *recovered* depending on $f_i(t_n)$ and $g_i(t_n)$. f_i describes the probability of infection regarding the misinformation, while g_i describes the probability of its debunking. They are defined in a way such that the number of infected and recovered neighbors $h_i^I(t_n)$ and $h_i^R(t_n)$ of node i influences its transition probabilities as soon as the information reaches a susceptible entity's neighbor. Both probabilities depend on the parameters α and β. While β determines the overall speed of the infection process, α describes the credibility of the information by weighting the number of infected neighbors, implying that the misinformation is more convincing than its true opposite. *Infected* entities may change their state to *recovered* if they receive the corresponding fact check given the probability p_{verify}. Recovered entities never change their state again. Thus, all entities will transition to the *recovered* state after a sufficient number of iterations.

Table 1: Summary of the model parameters

Notation	Description	Range
Social Network Parameters		
N	Number of nodes in the social network graph	$[1, \infty)$
k	Connectivity parameter of the BA graph	$[1, N)$
Information Diffusion Parameters		
α	Credibility of the information	$(0, 1)$
β	Spreading rate	$(0, 1)$
p_{verify}	Fact-checking probability	$(0, 1)$

3.3. Power Demand

The entities may change their power demand behavior depending on the information that they receive through social media. The change in power demand is modeled using a rule-based approach and depends only on the entities' states and the specific scenario. This part of the model is therefore decoupled from the social network and information propagation model. For a specific scenario under consideration, the rules need to be adjusted accordingly.

The model is based on several assumptions. Only *infected* entities are assumed to change their power demand, as only changes due to misinformation on social media are considered. Since *susceptible* entities are not aware of the misinformation and *recovered* entities do not believe in the information, these entities do not change their behavior and follow regular demand patterns.

Since misinformation affects people's behavior, which is mostly linked to power demand from household appliances rather than power demand from public or industrial buildings, we focus on domestic power demand only. Each node

in the social network graph represents a household, rather than an individual because household members can be considered densely connected and household appliances and EVs can be modeled per household.

The actual power demand is determined at appliance level. Infected households are assumed to consume additional power on top of their regular demand, as defined by the power consumption of the relevant appliances in each household given the specific scenario. Each type of appliance is distributed over all households with a specified adoption rate. If a household owns the appliance, it is assumed to be turned on when the household becomes "infected" and to be turned off after a specified duration. To determine the overall load on the power system, the aggregated power demand is calculated as the sum of each individual demand.

4. Parameter Estimation with Social Media Data

The model presented in the previous section can be used to simulate the effect of information propagation on power demand for different scenarios and parameters. However, there are several model parameters which have to be set. If the model shall create realistic progressions of information diffusion, the parameters have to be selected reasonably.

We aim to utilize social media data in order to obtain parameters such that the SIR model generates typical infection progressions, resembling propagation dynamics observed in real-world examples. Similar to other works [23, 36], we assume that during an extraordinary event, the spread of information is reflected in the use of social media, and that the number of social media posts related to the event is proportional to the number of infections in the model over time. While previous works that dealt with parameter estimation focused on differential equation based SIR models [23, 36, 37], we will connect this system-level view with the stochastic graph-based representation from our model to perform the estimation.

The parameter estimation is embedded in part 3 of the monitoring framework (see Figure 1). We differentiate between online and offline estimation. The online estimation would extrapolate the ongoing propagation process for the current situation. However, this would require to estimate model parameters during the information spread in real-time only from the collected social media posts, which is challenging because the simulation should be finished before the peak electricity demand is reached. Assuming universal model parameters enables the creation of a generalized information propagation model in advance. Due to the universal nature of information propagation on social media, regardless of the specific topic of the event, the model can be considered to generally represent propagation events once the parameters have been fitted. The parameters can be determined with social media data from past events, which is described in the following.

4.1. From Graph-based to System-level View

The SIR model used in this work is graph-based, i.e., the states of the nodes in the system are calculated individually and depend on the node's neighbors. In theory, the following optimization algorithm could be applied using the graph-based model, but the increased computational burden for its simulation and the stochastic nature of the model would complicate the estimation. Therefore, we transform the SIR model from the entity representation to the system-level view.

The system behavior of the SIR model is approximated using mean-field theory. First, the total number of entities in each state X at time step n can be written as $X_n = p_n^X \cdot N$. Thus, in average, $s_i(t_n) = \left[p_n^S, p_n^I, p_n^R\right]$. Second, we simplify the underlying graph structure by assuming that all nodes have the same mean degree $2k$ and the states are uniformly distributed over the network. Then, we can generalize the neighboring functions f_i, g_i (2) to (3) for the affected nodes by approximating the number of neighbors in state X as $h_i^X(t_n) = 2k \cdot p_n^X$:

$$f_n = c\beta \frac{(1+\alpha)I_n}{(1+\alpha)I_n + (1-\alpha)R_n} \tag{3a}$$

$$g_n = c\beta \frac{(1-\alpha)R_n}{(1+\alpha)I_n + (1-\alpha)R_n} \tag{3b}$$

The factor c accounts for the fact that not for all susceptible nodes $f_i(t_n) > 0$ and is set to $c = k/N$.

Consequently, the state equations (1) in time-discrete form can be simplified as the expected value of the number of entities in each state:

$$S_{n+1} = (1 - f_n - g_n)\, S_n \tag{4a}$$
$$I_{n+1} = f_n\, S_n + (1 - p_{\text{verify}})\, I_n \tag{4b}$$
$$R_{n+1} = g_n\, S_n + p_{\text{verify}}\, I_n + R_n \tag{4c}$$

4.2. Optimization Problem

The optimal model parameters, which fit the social media data best, can be determined by solving an optimization problem with the transformed SIR model (4). By extracting the social media posts related to an extraordinary event, the number of posts can be obtained as a function of time, denoted as $m_{\text{posts}}(t)$.

The model parameters are fitted to the data by minimizing the mean squared error between $m_{\text{posts}}(t)$ and the number of infected entities $I(t)$ computed by the system-level SIR model. The resulting problem can be represented as an optimization problem with non-linear difference equation constraints and is summarized as follows:

$$\min_{\alpha,\beta,p_{\text{verify}}} \sum_{n=0}^{n_{\text{end}}} \left(m_{\text{posts,n}} - I_n\right)^2 \tag{5a}$$
$$\text{s.t. } (4)\ \forall n \in \{0, \ldots, n_{\text{end}}\} \tag{5b}$$
$$S_0 = N - I_0 \tag{5c}$$
$$I_0 = m_{\text{posts,0}} \tag{5d}$$
$$R_0 = 0 \tag{5e}$$
$$\alpha, \beta, p_{\text{verify}} \in (0, 1) \tag{5f}$$

In the optimization process, the parameters $\alpha, \beta, p_{\text{verify}}$ are determined. The progression of S_n, I_n, R_n is calculated by solving the difference equations over all data points $m_{\text{posts,n}}$.

5. Case Study: False Pricing Attack on Demand Response

In this section, we apply the proposed parameter estimation algorithm of Section 4 with social media data to fit the parameters of the SIR model. With the refined information propagation and power demand model, presented in Section 3, we investigate a false pricing attack and the resulting effects on the power load. Additionally, we analyze the impact of EV charging on the results in the scenario. The evaluation of the model is concluded with a sensitivity analysis.

5.1. Scenario

We investigate the scenario of a false pricing attack in which a demand response (DR) event is falsely announced to incentivize consumers to increase their electricity demand at a given time. Malicious actors could create a fake social media post (see Figure 3), impersonating an energy supplier and announcing a limited-time offer for reduced electricity prices when the grid is at limit. According to a survey, it could be assumed that a significant percentage of customers would believe the information, eventually forward it, and change their electricity usage, regardless whether they are actually participating in a DR program [12].

Figure 3: Fictive post of a user impersonating an energy supplier

5.2. Social Media Dataset and Parameter Estimation

As discussed in Section 4, the use of real social media data is desirable to apply the parameter estimation algorithm for obtaining suitable parameters for the SIR model. As the method fits the data to the infection curve created by the SIR model, the data should cover a period where a certain, extraordinary event led to increased information propagation through social media. To capture only the dynamic of such an event, the social media posts may need to be filtered to the specific event.

For this work, we used tweets in relation to a forest fire caused by an explosion in the Grunewald forest during the night of August 3rd and 4th 2022, in Berlin, Germany [26]. 1759 tweets were collected between August 4, 2022 on 2:30 am and August 5, 2022 on 10:30 pm that contained the keyword "Grunewald" in combination with one or more terms from a predefined list of disaster-related terms, such as "fire", "explosion", or "bomb". The tweets were aggregated for every 30 minutes. Additionally, a moving average filter was applied to smooth the data.

The resulting time series of the number of posts $m_{\text{posts,n}}$ is used to estimate the parameters of the SIR model by solving the optimization problem introduced in (5). The numerical implementation is conducted with the Python package SciPy [38] by using the optimization method "Limited-memory BFGS". The algorithm yields the parameters $\alpha = 0.609$, $\beta = 0.006$, $p_{\text{verify}} = 0.036$.

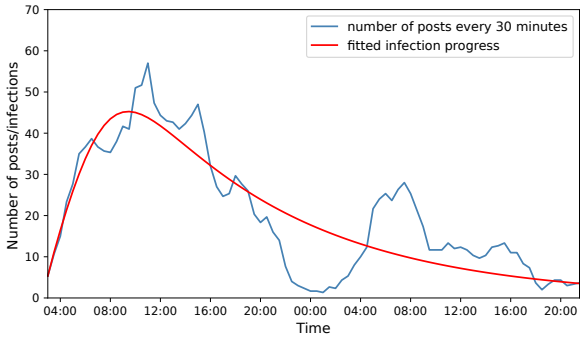
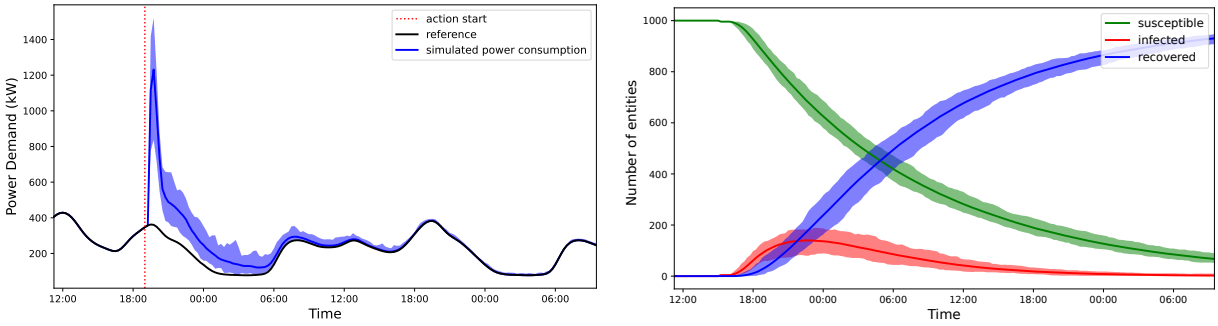


Figure 4: Grunewald dataset and the estimated infection progress

Figure 4 shows the number of posts from the filtered data and the infection process for the estimated parameters with the number of infected entities I_n over time. Beside apparent fluctuations in the use of social media, increased activity related to the event throughout the two days is visible. The resulting infection curve demonstrates that the estimated parameters can generate a similar infection process, especially for the first day of the event. The simulated infection process reaches its peak at a time which is almost identical to the peak observed in the data. Given that the event and the usage of social media both persist for more than a single day, a decline in activity during the night is observed, which cannot be represented by the model.

5.3. False Pricing Attack Scenario Simulation

With the estimated parameters from the previous subsection, we simulate the model for the scenario to analyze the altered power consumption and demand peaks.



(a) Regular and altered power demand

(b) Infection process

Figure 5: Simulation results for the demand response attack scenario

The dissemination of misinformation regarding reduced electricity prices is presumed to commence at 16:00, while the *infected* entities begin to act in accordance with the announced time (19:00), activating household appliances. The model parameters and appliances considered for this scenario are listed in Table 2. The adoption rate of the appliances is chosen as the percentage of German households that possess the specific items, as provided by the Federal Statistical Office of Germany [39]. The time step is set to 30 minutes. The regular power demand is modeled with standard load profiles given by the German Federal Association of Energy and Water Management (BDEW).

Table 2: Model parameters for the simulation

Social Network Parameters			
N	1000		
k	50		
Appliances			
Type	**Duration**	**Adoption Rate**	**Power Usage**
Washing machine	1.5 h	95 %	1 kW
Dryer	1 h	42.3 %	3 kW
Dishwasher	3 h	71.9 %	1.4 kW
Oven	1 h	80 %	2.6 kW
Electric stove	30 min	94 %	4 kW
EV charging	6 h	2 %	11 kW

Figure 5a shows the reference power demand and the simulated power consumption, altered due to the demand response attack. The announced time at which the event shall start is marked. The simulation is run 100 times with different seeds to assess the resulting output range of the stochastic model. The respective average, minimum, and maximum values of the simulation results are shown.

The power consumption rapidly increases after the beginning of the event. The extremely steep increase happens due to the fact that the spreading of the misinformation already started 3 hours earlier and a significant number of entities is infected at 19:00, as can be seen in Figure 5b, showing the infection process. The peak demand occurs at 20:00 and is 3.42 times higher (in average) than the regular power demand. The power demand decreases after about 2 hours, which can be explained by the duration that most household appliances run. Nevertheless, the power demand remains notably higher than the regular power demand for about 9 hours because of the delayed infections of other entities. After 24 hours, the infection wave is almost over and remaining susceptible entities become directly recovered.

Remarkably, the number of infected entities is low compared to the number of susceptible and recovered entities, with the maximum of 13.9 % (in average) at 22:30. Note, that the electricity load peak occurs 2.5 hours before the infection peak. The load peak was created by only 11.4 % (in average) coincidentally infected entities.

Beside the average values, the variation of the simulation runs can be analyzed. The peak load varies from 825 to 1510 kW, which equals 2.29 to 4.19 times the regular demand. Also, in the period of moderate increased demand from 2 to 9 hours after the event, a load range of 150-250 kW is visible, which refers to 25-40 % deviation from the average. The infection process shows some variance in the number of entities in each state over time, although the overall progress dynamic remains similar.

5.4. Impact of EV Charging

Next, the simulation model is applied to investigate the impact of increased EV adoption and home-charging, especially for the synchronized occurrence of such significant loads. To analyze this effect, the simulation of the demand response attack scenario was set up with different EV adoption rates, while keeping the other simulation parameters (see Table 2) unchanged.

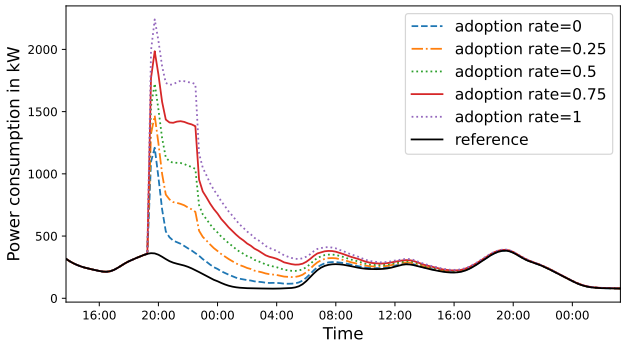


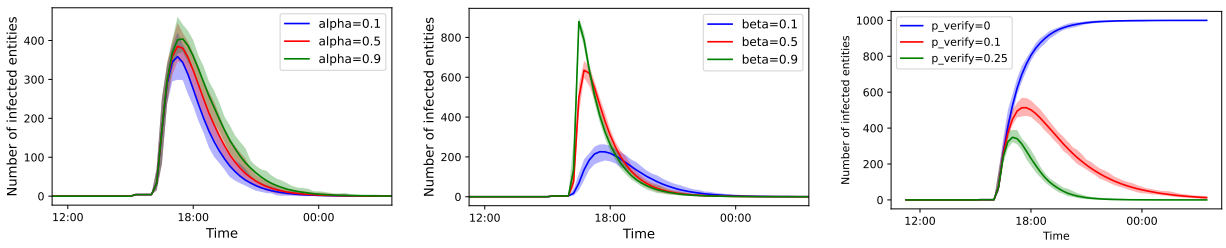
Figure 6: Simulated altered power demand with different EV adoption rates

Figure 6 shows the simulation results for the scenario with EV charging for adoption rates ranging from 0 to 100 %. Only the average results are shown here. The peak electricity demand increases by up to 85 % as the adoption rate increases. In addition, the duration of the abnormal load condition is longer than in the previous simulation, as EV charging is conducted for a longer period of time compared to household appliances.

5.5. Model Sensitivity Analysis

In contrast to the system-level SIR model, the graph-based SIR model employed in this work is based on stochastic equations (1). This implies that each simulation run can yield different results for the infection process. Incorporating randomness into a model is a viable approach to increase the variety of possible solutions, particularly in the case of uncertainty. However, it is essential to understand and quantify the variability of the stochastic model. Moreover, we want to analyze the influence of the parameters of the SIR model on the resulting infection process.

The sensitivity analysis is conducted by simulating the infection process by varying the model parameters $\alpha, \beta, p_{\text{verify}}$ separately while the other two parameters are fixed. The fixed values of the parameters are $\alpha = 0.4$, $\beta = 0.2$, $p_{\text{verify}} = 0.2$. To assess the variance of the results created by the randomness persistent in the SIR model, the simulation was run 100 times with different seeds for each parameter variation. The average, minimum, and maximum outcomes are determined.



(a) Infection process for varying α

(b) Infection process for varying β

(c) Infection process for varying p_{verify}

Figure 7: Results of the sensitivity analysis

A comparison of the results across different parameter variations in Figure 7 shows that the model seems to be insensitive to changes in α, while β and p_{verify} have a significant impact on the infection curve. A higher β leads to a higher and earlier peak of infections and faster decline, since β acts as a parameter for the speed of the infection process. Conversely, a higher p_{verify} leads to a smaller *and* shorter infection process because more entities recover faster. As the effect of p_{verify} comes into play at a later stage, the start of the infection curve is independent from its value.

Moreover, the number of entities over time is robust over multiple simulation runs. During the infection process, a certain amount of variance in the infection curve is visible.

6. Discussion

Although the investigated model includes multiple assumptions, the results of the case study demonstrate that a DR attack could in general create a sudden and significant impact on the aggregated power demand. The synchronized appliance usage triggered by a maliciously announced time window, combined with the high power consumption of "infected" entities, could push the aggregated load to critical and unprecedented levels. This poses a serious threat to power systems in three distinct ways.

- First, the significant power increase can exceed limits in distribution grids, leading to equipment damage, tripping, and local blackouts. If the scale of the event is large, it could also exceed limits in transmission grids and lead to load shedding.
- Second, the steep and sudden rise in demand within a short period could become critical for the frequency control, especially if this occurs unexpectedly, thus endangering grid stability. If the resulting load spike is high and rapid, in combination with the unexpected occurrence, a lack of generation reserve may result in voltage instability, load shedding, or rolling blackouts. A sudden *decrease* could be critical for grid stability as well [3].
- Third, the positive intention of DR can be subverted to detrimental effects when maliciously announced, potentially endangering power supply during periods of low energy availability.

The case study further demonstrates that the increasing adoption of EV home-charging severely increases the impact of a DR attack. Because of the longer time of usage, here it is not only about reaching load capacity limits, but also involves a significant amount of energy, depending on the affected scale. With the increasing electrification and the trend to an "all-electric" society, the effects of such events might increase in future and become more relevant.

Given the potential consequences and the rising vulnerability to synchronized consumer behavior facilitated by social media, such scenarios require consideration by policy makers for analyzing potential consequences and handling such situations effectively. Besides DR attacks, related scenarios involving information propagation and altered power demands could yield similar results. While grid capacity enhancements could avoid reaching limits, this solution is costly and the challenge of grid balancing remains. As the outcomes can usually be mitigated with sufficient reaction time, the early detection and load prediction for such anomalous events seems of vital importance.

Comparing with studies of similar scope, the results of our case study are qualitatively similar to other works that investigated DR attacks. The differences, for example the higher load peak compared to [11], can mostly be attributed to different participation rates and appliance parameters. While we considered solely the effect on the aggregated load over time without making assumptions on load limits, some works used a simulated power network model, either general reference grids [33, 34, 10] or artificially reconstructed grids [12, 25]. In that case, the failure of certain elements of a specific network can be assessed, under the assumption of simplified failure limits [33, 10, 12]. However, it is an open question, to which extent a network simulation would provide a benefit for non-localized demand trends. Given the several uncertainties involved in the considered extraordinary events, the aggregated load could be sufficient to assess whether the impact of a scenario is principally critical. Certainly, if detailed geo-referenced grid data are available and capacity limits are known, the precision of assessments and analysis of spatial effects for the altered power demand could be improved for a given case. Nevertheless, the challenge of mapping entities in the social network geographically to the electrical grid remains.

Regarding the modeling of the information propagation, the graph-based SIR model used in our study can produce various, robust infection progressions. In contrast to independent cascade models [33, 34, 12] and threshold-based influence models [10, 12] used in previous works, the SIR model includes a recovery state, which represents the debunking of misinformation, leading to the decrease in infections. Although the recovery state mostly comes into play at a later stage, it was shown that the fact checking probability has a significant influence on the infection peak (see Figure 7c). In addition, the SIR model can be easily extended with additional states to further refine the modeling of rumor spreading [22]. As the considered scenarios are subject to uncertainties, stochastic modeling is essential to provide a range of possible outcomes. While the independent cascade and threshold-based influence models are

graph-based and stochastic models, the SIR model used in [25] is based on deterministic differential equations. The developed SIR model of our study is graph-based and stochastic, thus can represent uncertainty and account for different network structures.

Selecting the parameters of the propagation models is a crucial aspect. We developed an optimization algorithm to fit the SIR model with social media data (see Section 4). In this regard, choosing the SIR model provides a further benefit because we can connect the graph-based SIR model with the system dynamics representation, simplifying the fitting from a top-down perspective. The case study showed that the parameter estimation can be used to obtain a fitted model that can reproduce propagation dynamics observed in real data. Most of the other studies did not use any data to estimate model parameters and displayed their results as a function of them [11, 12, 25]. Raman et al. [12] conducted a survey for estimating the propensity of people to believe and forward misinformation related to a DR attack, although the authors mentioned that the direct adoption of survey results into model parameters is challenging. While this can be seen as fitting on the low level, our model fitting takes place at the level of the aggregated behavior. This has the advantage that it ensures that the resulting model behavior is aligned with the real data.

Many of the studies considered the optimization of load shedding as a counter measure [33, 34, 10, 25]. As has emerged from the beginning of this section, the unexpectedness of the synchronized demand changes plays a central role in the criticality and the potential consequences. For this reason, the continuous monitoring of social media and online trends, as suggested in Section 1.1, could provide operators with timely warnings and predictions. This would increase the time window for effectively dealing with such critical situations, which would ultimately improve the chance of avoiding load shedding. To implement the real-time monitoring, multiple challenges remain. For example, in our model all entities can get in contact with misinformation. However, as not everyone might use social media and receive the information in the relevant time period, the number of entities, i.e. N, would need to be inferred from the amount of social media activity. Moreover, the SIR propagation parameters might be situation-specific. While we showed that the offline estimation can be used to capture the patterns of a social media dataset, the real-time estimation of model parameters during the propagation process could improve the prediction of the model.

7. Conclusion and Future Work

In this paper, we investigated the modeling and detection of social media induced impacts on domestic power demand. We presented a graph-based model for the propagation of information over social networks based on the epidemiological SIR model. Using the scenario of a misinformation attack, we demonstrated that sudden, unexpected demand actions can lead to significant over-consumption and pose a threat for grid stability. The increasing adoption of energy-intensive technologies in households, e.g. EV home-charging, amplifies the criticality of these scenarios. This shows that more research for understanding, detecting, and handling such power demand synchronization events is needed.

We presented the novel idea of a monitoring framework, providing a basis for analyzing the impact of potential scenarios and predicting critical load shifts in real-time. We used social media data from a past crisis event to estimate information propagation patterns for obtaining a realistic model. This approach extends current research, as mostly *reactive* countermeasures, such as load shedding, are considered. A paradigm shift to *proactive* measures could mitigate the negative consequences to power grid stability and prevent blackouts. Grid operators could use the monitoring framework to anticipate critical power demand peaks in advance, allowing them to take effective countermeasures, thus enhancing grid stability and resilience.

Future research could focus on the parts of the monitoring framework which were not considered in detail in this work. The collection, filtering, and processing of social media data involves several conceptual and practical challenges that require further investigation. In particular, the semantic information of social media posts and the potential effect on power demand needs to be extracted. Advanced natural language processing techniques or large language models could provide a solution for this task. For example, the language model CitEnergy [40], specifically designed to analyze energy-related tweets, could be employed. Furthermore, methods to map social media posts to geographic locations could be investigated to identify localized events [41]. With spatial information, social networks could be mapped more accurately to the power grid infrastructure to assess potential impacts with a higher precision.

CRediT authorship contribution statement

Isabella N. Grieser: Conceptualization, Methodology, Software, Writing - Original Draft. **Tobias Gebhard:** Conceptualization, Formal analysis, Methodology, Writing - Review & Editing. **Andrea Tundis:** Writing - Review & Editing, Project administration. **Jens Kersten:** Resources, Writing - Review & Editing. **Tobias Elßner:** Resources, Data curation, Writing - Review & Editing. **Florian Steinke:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and Code Availability

Data will be made available on request. The code developed for this study is available in the following GitHub repository: https://github.com/isabella-grieser/effects-of-misinformation-on-the-electrical-grid

Acknowledgment

This work has been conducted in the context of the "urbanModel" and "WebData4urbanModel" projects funded from the German Aerospace Center (DLR). It has been performed in the context of the LOEWE center emergenCITY [LOEWE/1/12/519/03/05.001(0016)/72].

References

[1] I. B. Sperstad, G. H. Kjølle, O. Gjerde, A comprehensive framework for vulnerability analysis of extraordinary events in power systems, Reliability Engineering & System Safety 196 (2020) 106788. doi:10.1016/j.ress.2019.106788.

[2] D. L. Donaldson, M. S. Alvarez-Alvarado, D. Jayaweera, Power system resiliency during wildfires under increasing penetration of electric vehicles, in: 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), IEEE, 2020. doi:10.1109/pmaps47429.2020.9183683.

[3] Can you have a big 'switch off'?, BBC, (accessed 2024-08-20) (2007).
URL http://news.bbc.co.uk/2/hi/uk_news/magazine/6981356.stm

[4] Y. Lei, Y. Zhang, J. Guo, D. Zhou, J. Culliss, P. Irminger, Y. Liu, The impact of synchronized human activities on power system frequency, in: 2014 IEEE PES General Meeting | Conference & Exposition, 2014, pp. 1–5, ISSN: 1932-5517. doi:10.1109/PESGM.2014.6939472.

[5] M. Naeem, Do social media platforms develop consumer panic buying during the fear of covid-19 pandemic, Journal of Retailing and Consumer Services 58 (2021) 102226. doi:10.1016/j.jretconser.2020.102226.

[6] Coronavirus crisis: Experimental data reflect buying behaviour in retail trade, Federal Statistical Office, (accessed 2024-08-20) (2020).
URL https://www.destatis.de/EN/Press/2020/03/PE20_112_61.html?nn=395260

[7] T. Althoff, P. Jindal, J. Leskovec, Online actions with offline impact: How online social networks influence online and offline user behavior, in: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, WSDM 2017, ACM, 2017. doi:10.1145/3018661.3018672.

[8] U. K. H. Ecker, S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, M. A. Amazeen, The psychological drivers of misinformation belief and its resistance to correction, Nature Reviews Psychology 1 (1) (2022) 13–29. doi:10.1038/s44159-021-00006-y.

[9] E. M. Wells, M. Boden, I. Tseytlin, I. Linkov, Modeling critical infrastructure resilience under compounding threats: A systematic literature review, Progress in Disaster Science 15 (2022) 100244. doi:10.1016/j.pdisas.2022.100244.

[10] D. Tang, Y.-P. Fang, E. Zio, J. E. Ramirez-Marquez, Resilience of smart power grids to false pricing attacks in the social network, IEEE Access 7 (2019) 80491–80505. doi:10.1109/access.2019.2923578.

[11] G. Raman, J. C.-H. Peng, T. Rahwan, Manipulating residents' behavior to attack the urban power distribution system, IEEE Transactions on Industrial Informatics 15 (10) (2019) 5575–5587. doi:10.1109/tii.2019.2903882.

[12] G. Raman, B. AlShebli, M. Waniek, T. Rahwan, J. C.-H. Peng, How weaponizing disinformation can bring down a city's power grid, PLOS ONE 15 (8) (2020) e0236517. doi:10.1371/journal.pone.0236517.

[13] N. LaLone, A. Tapia, C. Zobel, C. Caraega, V. K. Neppalli, S. Halse, Embracing human noise as resilience indicator: twitter as power grid correlate, Sustainable and Resilient Infrastructure 2 (4) (2017) 169–178. doi:10.1080/23789689.2017.1328920.

[14] S. Jamalzadeh, L. Mettenbrink, K. Barker, A. D. González, S. Radhakrishnan, J. Johansson, E. Bessarabova, Weaponized disinformation spread and its impact on multi-commodity critical infrastructure networks, Reliability Engineering & System Safety 243 (2024) 109819. doi:10.1016/j.ress.2023.109819.

[15] T. Gebhard, E. Brucherseifer, F. Steinke, Monitoring electricity demand synchronization using copulas, in: Proc. IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2022. doi:10.1109/ISGT-Europe54678.2022.9960369.

[16] C. Kuster, Y. Rezgui, M. Mourshed, Electrical load forecasting models: A critical systematic review, Sustainable Cities and Society 35 (2017) 257–270. doi:10.1016/j.scs.2017.08.009.

[17] N. Fumo, M. A. Rafe Biswas, Regression analysis for prediction of residential energy consumption, Renewable and Sustainable Energy Reviews 47 (2015) 332–343. doi:10.1016/j.rser.2015.03.035.

[18] T. Bodnar, M. L. Dering, C. Tucker, K. M. edu, Using large-scale social media networks as a scalable sensing system for modeling real-time energy utilization patterns, IEEE Transactions on Systems, Man, and Cybernetics: Systems 47 (10) (2017) 2627–2640. doi:10.1109/tsmc.2016.2618860.

[19] J. Heglund, K. M. Hopkinson, H. T. Tran, Social sensing: towards social media as a sensor for resilience in power systems and other critical infrastructures, Sustainable and Resilient Infrastructure 6 (1–2) (2020) 94–106. doi:10.1080/23789689.2020.1719728.

[20] C. Deng, W. Lin, X. Ye, Z. Li, Z. Zhang, G. Xu, Social media data as a proxy for hourly fine-scale electric power consumption estimation, Environment and Planning A: Economy and Space 50 (8) (2018) 1553–1557. doi:10.1177/0308518x18786250.

[21] X. Wu, C. Dou, D. Yue, Electricity load forecast considering search engine indices, Electric Power Systems Research 199 (2021) 107398. doi:10.1016/j.epsr.2021.107398.

[22] S. Raponi, Z. Khalifa, G. Oligeri, R. Di Pietro, Fake news propagation: A review of epidemic models, datasets, and insights, ACM Transactions on the Web 16 (3) (2022) 1–34. doi:10.1145/3522756.

[23] F. Jin, E. Dougherty, P. Saraf, Y. Cao, N. Ramakrishnan, Epidemiological modeling of news and rumors on twitter, in: Proceedings of the 7th Workshop on Social Network Mining and Analysis, KDD' 13, ACM, 2013. doi:10.1145/2501025.2501027.

[24] M. Tambuscio, G. Ruffo, A. Flammini, F. Menczer, Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks, in: Proceedings of the 24th International Conference on World Wide Web, WWW '15, ACM, 2015. doi:10.1145/2740908.2742572.

[25] S. Jamalzadeh, K. Barker, A. D. González, S. Radhakrishnan, Protecting infrastructure performance from disinformation attacks, Scientific Reports 12 (1). doi:10.1038/s41598-022-16832-w.

[26] Series of explosions hamper attempts to tackle berlin forest fire, The Guardian, (accessed 2024-08-20) (2022).
URL https://www.theguardian.com/world/2022/aug/04/series-of-explosions-hamper-attempts-to-tackle-berlin-forest-fire

[27] G. Timilsina, J. Steinbuks, Economic costs of electricity load shedding in nepal, Renewable and Sustainable Energy Reviews 146 (2021) 111112. doi:10.1016/j.rser.2021.111112.

[28] M. Schneider, O. H. Ramírez-Agudelo, L. Halekotte, D. Lichte, A probabilistic approach to dynamic risk scenario identification, in: 32nd European Safety and Reliability Conference, ESREL, Research Publishing Services, 2022. doi:10.3850/978-981-18-5183-4_r06-01-282-cd.

[29] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, B. Bhattacharjee, Measurement and analysis of online social networks, in: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, IMC07, ACM, 2007. doi:10.1145/1298306.1298311.

[30] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (6684) (1998) 440–442. doi:10.1038/30918.

[31] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512. doi:10.1126/science.286.5439.509.

[32] T. Britton, Stochastic epidemic models: A survey, Mathematical Biosciences 225 (1) (2010) 24–35. doi:https://doi.org/10.1016/j.mbs.2010.01.006.

[33] T. Pan, S. Mishra, L. N. Nguyen, G. Lee, J. Kang, J. Seo, M. T. Thai, Threat from being social: Vulnerability analysis of social network coupled smart grid, IEEE Access 5 (2017) 16774–16783. doi:10.1109/access.2017.2738565.

[34] L. N. Nguyen, J. D. Smith, M. T. Thai, Vulnerability assessment of social-smart grids: An algorithmic approach, in: 2019 IEEE Global Communications Conference (GLOBECOM), 2019, pp. 1–7, ISSN: 2576-6813. doi:10.1109/GLOBECOM38437.2019.9014195.

[35] Q. Zhang, A. Boukerche, A novel infrastructure-based worm spreading countermeasure for vehicular networks, IEEE Transactions on Intelligent Transportation Systems 19 (7) (2018) 2188–2203. doi:10.1109/tits.2018.2807358.

[36] M. Maleki, E. Mead, M. Arani, N. Agarwal, Using an epidemiological model to study the spread of misinformation during the black lives matter movement (2021). doi:10.48550/ARXIV.2103.12191.

[37] M. Castiello, D. Conte, S. Iscaro, Using epidemiological models to predict the spread of information on twitter, Algorithms 16 (8) (2023) 391. doi:10.3390/a16080391.

[38] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.

[39] Equipment of households with electrical household appliances and others (germany), Federal Statistical Office, (accessed 2024-08-20) (2018).
URL https://www.destatis.de/EN/Themes/Society-Environment/Income-Consumption-Living-Conditions/Equipment-Consumer-Durables/Tables/liste-equipment-households-electrical--household-appliance-others-germany.html

[40] J. Bedi, D. Toshniwal, Citenergy: A bert based model to analyse citizens' energy-tweets, Sustainable Cities and Society 80 (2022) 103706. doi:10.1016/j.scs.2022.103706.

[41] X. Hu, Z. Zhou, H. Li, Y. Hu, F. Gu, J. Kersten, H. Fan, F. Klan, Location reference recognition from texts: A survey and comparison, ACM Comput. Surv. 56 (5). doi:10.1145/3625819.