

Infrastructure System Performance under Weaponized Disinformation

Kash Barker, Ph.D.

Industrial and Systems Engineering
University of Oklahoma



Background

Recent work

Our activities

Concluding remarks



Disinformation

Defined as “information falsely characterizing the state of the system, including rumors, factual errors, and attempts at deception”

[Floridi 2005]

Disinformation is indeed rising on online platforms

[Vosough et al. 2018, Alcott et al. 2019]



“

“No technology has been weaponized
at such an unprecedented global
scale as social media.”

- Ong and Cabañes, from *Architects of Networked
Disinformation*, 2018



Recent examples

“86% of online global citizens
tend to be at least initially fooled
by fake news stories”


[Simpson 2019]



INDEPENDENT

Many trust what their friends post on social media more than experts, poll claims

Researchers warn of the dangers of ruling out the cleverest as they can assist in helping us understand issues

Monday 22 July 2019 15:07 BST •  Comments



Exploiting infrastructure systems



The fragility and vulnerability of infrastructure systems have exposed society to risks of disruption with severe consequences

Direct attacks on cyber-driven infrastructure systems have been well-studied for years

- Viruses
- Ransomware
- False data injections

Exploiting infrastructure systems

However, an emerging, over-the-horizon threat: **adversaries who attack infrastructure systems indirectly through disinformation**

Altering consumption behavior of unwitting users influenced by weaponized disinformation shared through social networks

Perhaps?

...An airline
passenger posts
on social media
a false alert of a
suspicious
package in a
restroom at
O'Hare airport

Those who monitor social media
immediately send emergency signals
Personnel move passengers from the
airport, cancel outgoing flights,
prevent incoming flights from
landing

A similar event occurred at London's
Gatwick airport [Thomaselli 2020]

Perhaps?

...A social media user falsely reports a major accident on a busy segment of I-695 near Washington DC

Algorithms that compute shortest paths use this information and update Uber and Lyft routes
This causes congestion on new routes, freeing up old routes, which in turn creates another update in routes, resulting in oscillation effects and erratic driving behaviors

Similar events occurred globally
[Tufnell 2014, Barrett 2020]

Perhaps?



A look at our no-good, dirty, rotten record-breaking February weather

by Christopher Nestman, KTUL staff | Mon, March 1st 2021, 12:48 PM CST



National Centers for
Environmental Information
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

The Great Texas Freeze: February 11–20, 2021

Background

Recent work

Our activities

Concluding remarks



Recent work

Renewable energy sources (e.g., solar, wind) can be variable in their ability to meet demand
Backup sources tend to be high-emission gas-turbine-based power plants

Consumer-centric demand response services can help manage this variability

Messages to consumers to alter demand in real-time

Usually textual (e.g., text message, email, social media)



Recent work



IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 15, NO. 10, OCTOBER 2019

5575

Manipulating Residents' Behavior to Attack the Urban Power Distribution System

Gururaghav Raman , *Student Member, IEEE*, Jimmy Chih-Hsien Peng , *Member, IEEE*,
and Talal Rahwan

Explored user behavior models, probability of adoption of fake demand response messages

Integrated with a power grid model to understand impact on a city's power grid

Incorporated defender actions to counter disinformation

Recent work

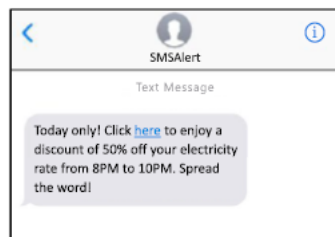
PLOS ONE

How weaponizing disinformation can bring down a city's power grid

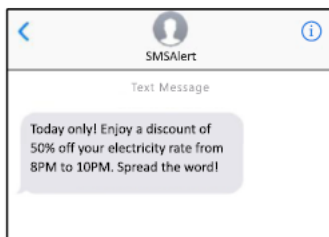
Gururaghav Raman, Bedoor AIShebli , Marcin Waniek , Talal Rahwan , Jimmy Chih-Hsien Peng 

Published: August 12, 2020 • <https://doi.org/10.1371/journal.pone.0236517>

Explore different messages that might influence power consumption behavior: what was the message? who did it come from?



Message 1



Message 2



Message 3



Message 4

Recent work

PLOS ONE

How weaponizing disinformation can bring down a city's power grid

Gururaghav Raman, Bedoor AlShebli , Marcin Waniek , Talal Rahwan , Jimmy Chih-Hsien Peng 

Published: August 12, 2020 • <https://doi.org/10.1371/journal.pone.0236517>

Considers EV adoption rates, as EVs: (i) increases vulnerability due to consumers controlling more deferrable energy, and (ii) increases resilience due to the grid's upgraded capacity

An influence model (cascading + linear threshold) determined adoption, overloading, and removal of lines in a power simulation

Recent work

IEEE Access

Resilience of Smart Power Grids to False Pricing Attacks in the Social Network

DAOGUI TANG¹, YI-PING FANG¹, (Member, IEEE), ENRICO ZIO^{2,3}, (Senior Member, IEEE),
AND JOSE EMMANUEL RAMIREZ-MARQUEZ⁴

Incorporated personality traits into an influence model
Related a social network to an underlying power grid
Simulated several iterations of influence spread based on uncertain user behaviors

Recent work

scientific reports
Traffic networks are vulnerable to disinformation attacks

[Marcin Waniek](#), [Gururaghav Raman](#), [Bedoor AlShebli](#), [Jimmy Chih-Hsien Peng](#) ✉ & [Talal Rahwan](#) ✉

[Scientific Reports](#) **11**, Article number: 5329 (2021) | [Cite this article](#)

Adversary spreads false traffic alerts (e.g., accident, congestion) with the aim of manipulating the routes taken by drivers in a city

Convergence vs divergence messages

Survey of user behavior on likelihood to follow-through on a notification

Integration with a traffic simulation, additional trip time measured

Recent work



Annals of Operations Research

Editorial | [Published: 04 July 2023](#)

Editorial: fake news, misinformation, and supply chain disruptions: the role of emerging technologies

[Konstantina Spanaki](#) , [Thanos Papadopoulos](#), [Uchitha Jayawickrama](#), [Femi Olan](#) & [Shaofeng Liu](#)

Special issue with mostly qualitative ideas, some survey-driven
Several applications in COVID-19 misinformation

Background

Recent work

Our activities

Concluding remarks



Weaponized disinformation

We conceptualize

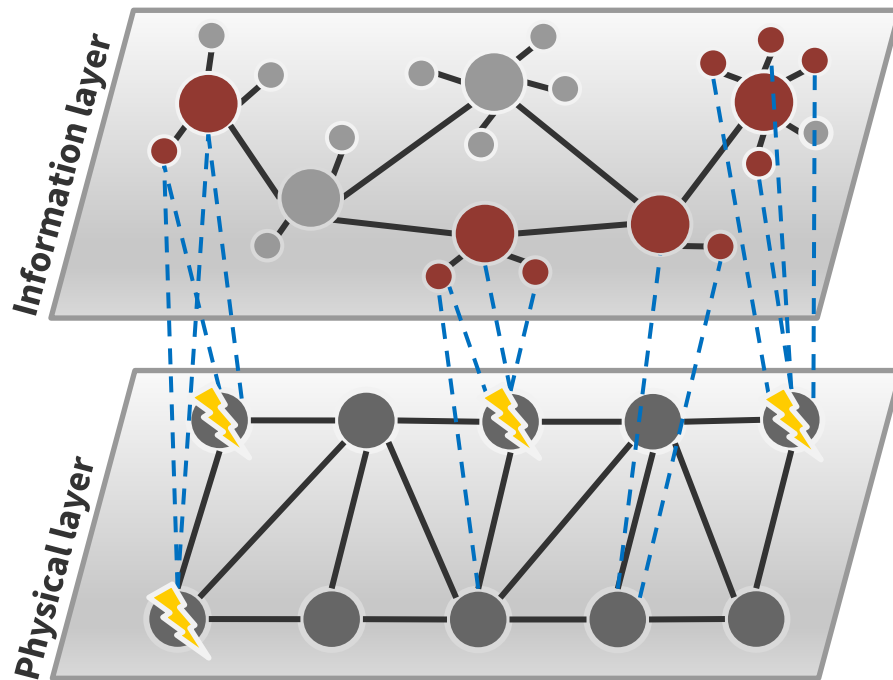
An information layer

A physical layer

Connections in between
driven by human behavior



Saeed Jamalzadeh, PhD
Bayer



Weaponized disinformation

Some recent work

Epidemiological model:

susceptible-infected-recovered

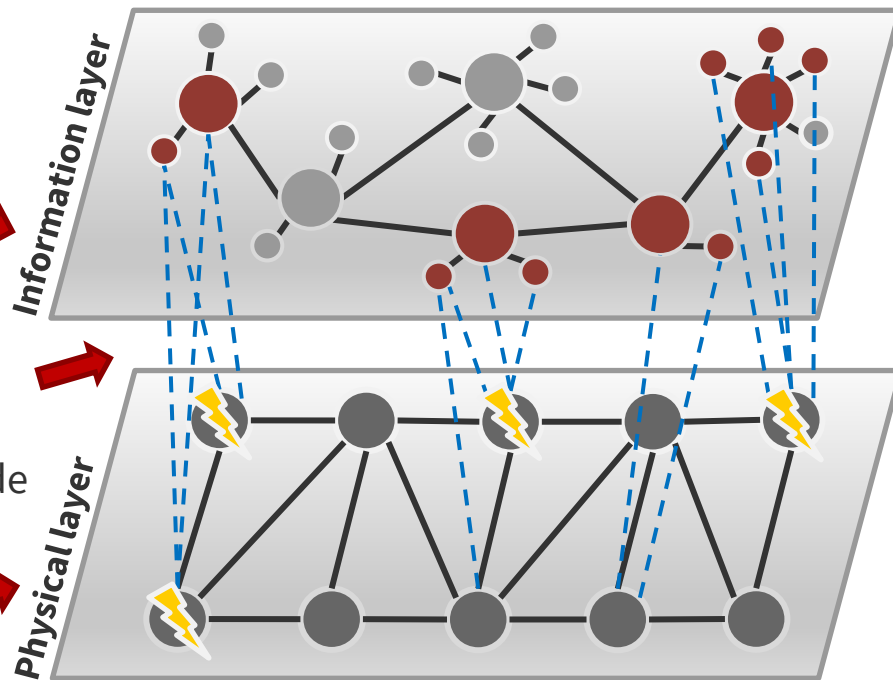
Terms governing disinformation spread (β), recovery (γ)

Simple threshold

Some number of users receive/act on disinfo in the vicinity of a physical node

Optimization model: network flows

Decision variable (g_{it}) governing spread of good information at node i , time t



Weaponized disinformation

Multi-commodity extension



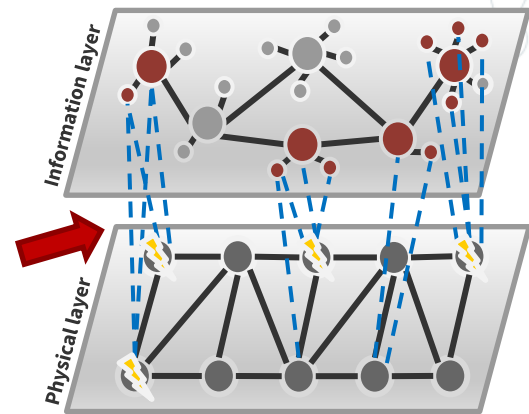
Lily Mettenbrink, MS
Mitre



Relationships between layers



Alice Nyanzi
OU ISE PhD student



Weaponized disinformation

Output

Jamalzadeh, S., K. Barker, A.D. González, S. Radhakrishnan. 2022. INITIAL MODEL, ELECTRIC POWER GRID. *Scientific Reports*, 12: 12707.

Jamalzadeh, S., L. Mettenbrink, K. Barker, A.D. González, S. Radhakrishnan, J. Johansson. 2024. MULTI-COMMODITY EXTENSION, *Reliability Engineering and System Safety*, 243: 109819.

Output

Barker, K., et al. 2024. PERSPECTIVES PAPER. In second revision in *Risk Analysis*.

Jamalzadeh, S., K. Barker, A.D. González, S. Radhakrishnan, G. Sansavini. 2025. INFORMATION INTERDICTION, NY SUBWAY. *Journal of Complex Networks*, 2: cnaf003.

Weaponized disinformation

Output

Jamalzadeh, S., K. Barker, A.D. González, S. Radhakrishnan, and E. Bessarabova. 2025. ROBUST FORMULATION. *Physica A*, 660: 130365.

Nanyanzi, A., K. Barker, S. Radhakrishnan, Z. Zhang, and J.E. Ramirez-Marquez. 2024. CAUSAL FAILURES. Submitted to *Reliability Engineering and System Safety*.

Output

Khameneh, R.T., J.E. Ramirez-Marquez, and K. Barker. 2025. DISINFO + PUBLIC TRANSIT. *Reliability Engineering and System Safety*, 255: 110656.

Rocco, C.M., K. Barker, S. Radhakrishnan, and J.E. Ramirez-Marquez. 2025. MULTI-OBJECTIVE INTERDICTION. *Social Network Analysis and Mining*, 15: 28.

Weaponized disinformation + transit

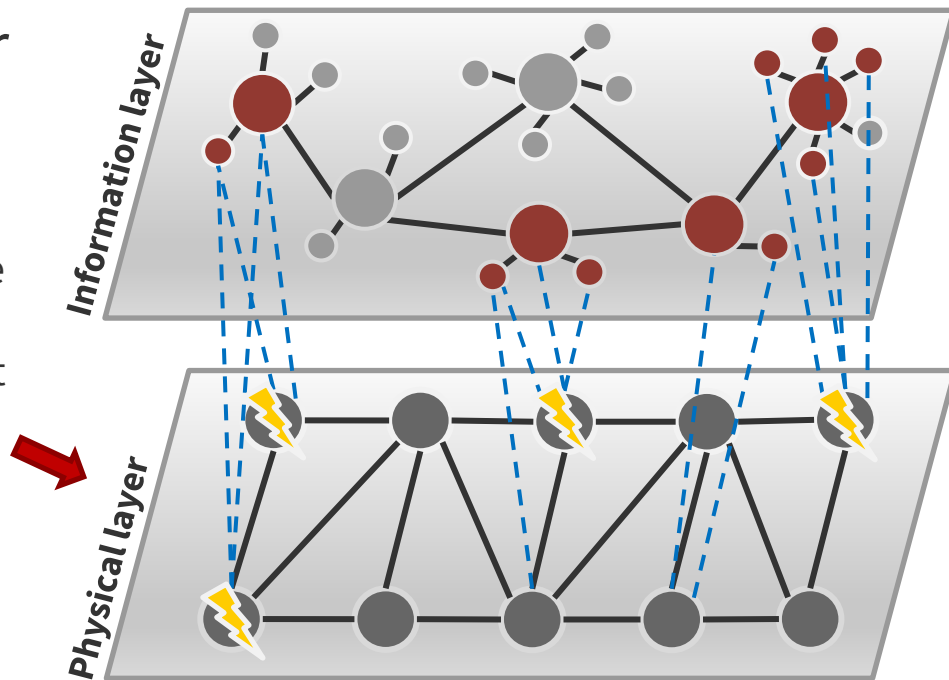
Let's look at a particular problem

Public transit network

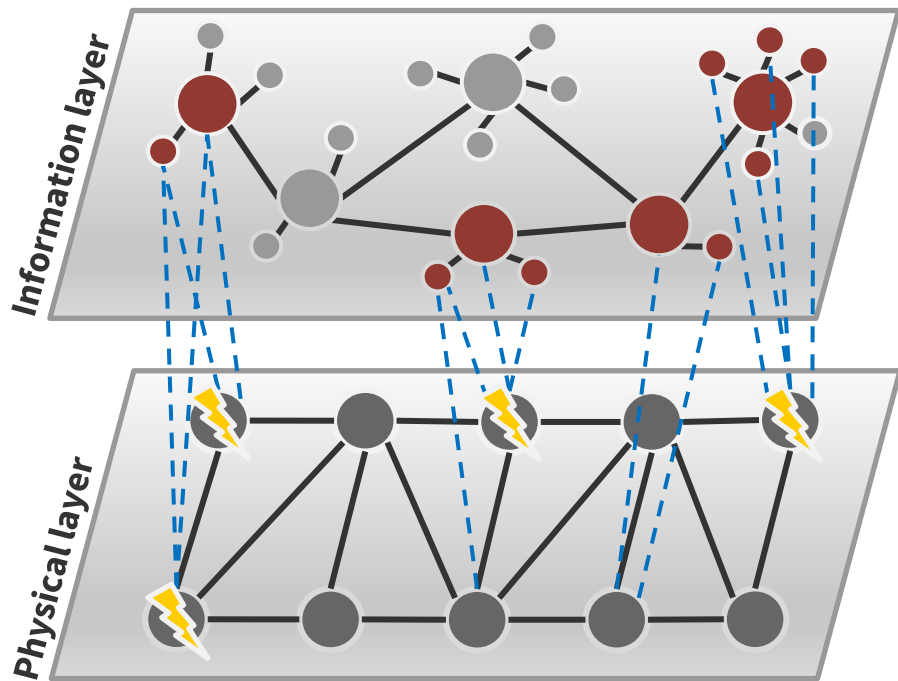
Users are informed or disinformation, where they believe certain metro stations are non-operational and thus find routes that avoid those stations or they use another mode (e.g., walking)



Daniel Cabrera Giraldo
OU ISE PhD student



Weaponized disinformation + transit



Information layer: $G(U, L)$

$U_s \subseteq U$, susceptible to disinfo

$U_i \subseteq U_s$, susceptible transit users

$U_b \subseteq U$, disinfo spreaders

$U_c \subseteq U$, correct info spreaders

Physical layer: $G(N, A)$

N , transit stations

A , transit segments

Weaponized disinformation + transit

- ◎ MILP model integrates user behavior with physical transportation flows
- ◎ We account for behavioral responses to false information and optimally select users to disseminate accurate updates

$$\min \sum_{a \in A, k \in K, t \in T} c_a f_{a,k,t}$$

Minimize the aggregate travel time of all users, including disinformed and informed users

Weaponized disinformation + transit

s. t.

$$\sum d_{u,u'} \geq y_u \quad \forall u \in U_i$$

Determines that y_u , which represents the proportion of links that spread disinfo to user u out of all links (u, u') connected to u , can only be >0 when u has at least one connection that spreads disinfo

$$\sum d_{u,u'} z_{u,u'} + (y_u - 1) \sum d_{u,u'} + y_u \sum d_{u,u'} = 0 \quad \forall u \in U_i$$

Determines the value of y_u depending on the direct connections of u if they are susceptible

Weaponized disinformation + transit

s. t.

$$y_u - z_{u,u'} \geq 0 \quad \forall u \in U_i, u' \in U_s: (u, u') \in L_s$$

$$x_u - z_{u,u'} \geq 0 \quad \forall u \in U_i, u' \in U_s: (u, u') \in L_s$$

$$1 + z_{u,u'} - x_{u'} - y_u \geq 0 \quad \forall u \in U_i, u' \in U_s: (u, u') \in L_s$$

Governs the protected or disinformed status of susceptible users

$$\sum x_u \leq n$$

Limits the number of users selected to spread correct information, optimizing resource allocation for intervention

$$\phi - \pi_u y_u + p_u \geq 0$$

$$\phi - \pi_u y_u + p_u \leq 1$$

Determine if a user will be disinformed based on their proportion of disinformed links

Weaponized disinformation + transit

s. t.

$$\sum f_{(i,j,w),k,t} - \sum f_{(j,i,w),k,t} = \sum g_{(u,k,i)} \quad \forall k \in K, i \in N$$

Maintains flow conservation at each station, ensuring total inflow and outflow are equal to the demand at that node for users of a specific OD pair

$$\sum f_{(i,j,w),k,t} - \sum f_{(j,i,w),k,t} = \sum p_u g_{(u,k,i)} \quad \forall k \in K, i \in N: t = DI$$

Matches disinformed (DI) users in the information layer with their counterparts in the physical layer, so that they avoid specific routes

Weaponized disinformation + transit

s. t.

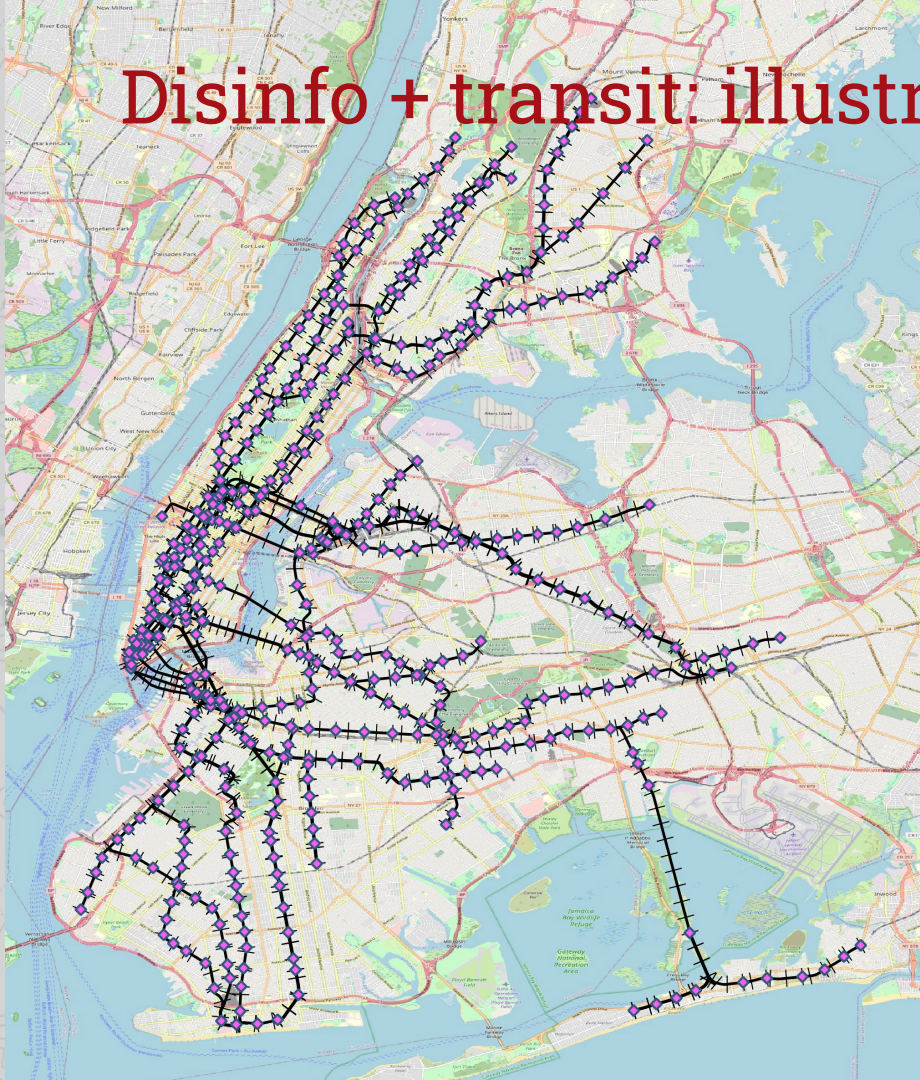
$$\sum f_{a,k,t} = 0 \quad \forall a \in A^0: t = DI$$

Prohibits disinformed users from traversing links associated with stations reported to be disrupted, either reroute through alternative paths that avoid the reportedly disrupted areas or switch to a different mode

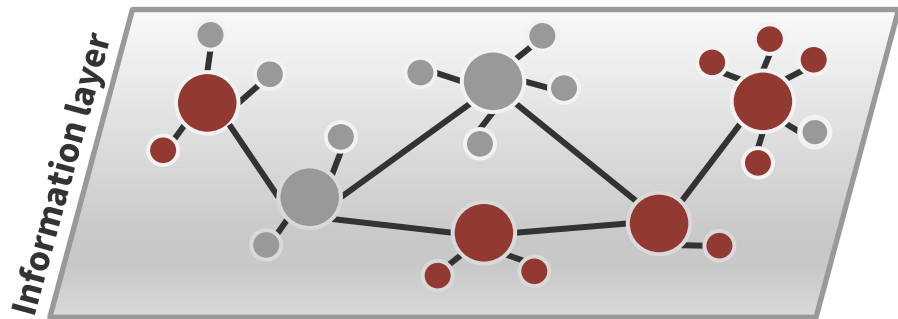
$$\sum f_{a,k,t} \geq 0 \quad \forall a \in A, k \in K, t \in T$$

Ensures that flow across any link remains positive

Disinfo + transit: illustrative example



Disinfo + transit: illustrative example

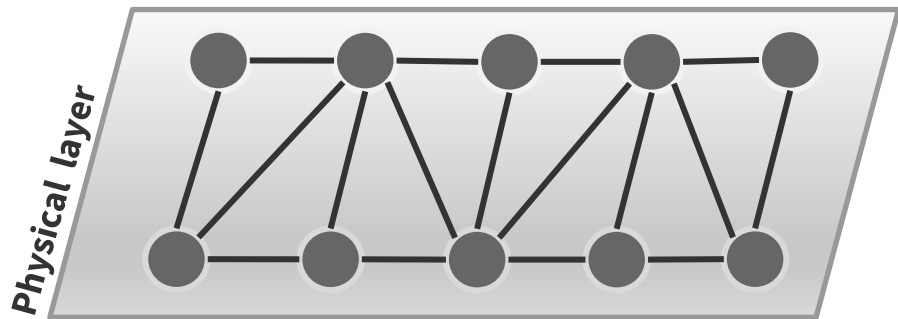


Constructed from a public Twitter graph (80K nodes, 1.7M links), randomly sampling 20,000 users, retaining their mutual interactions

Assigned each user a susceptibility score

Used studies of percentages that believe/share disinfo: 15% have shared disinfo knowingly, assumed 60% of susceptible users actively engage with the subway system

Disinfo + transit: illustrative example



Constructed using OD passenger flow estimates from a Bayesian inference approach modeling passengers traveling between stations

Limited the transit network (205 nodes, 680 links) to the seven principal subway lines operating during the busiest hour of the day due to computation

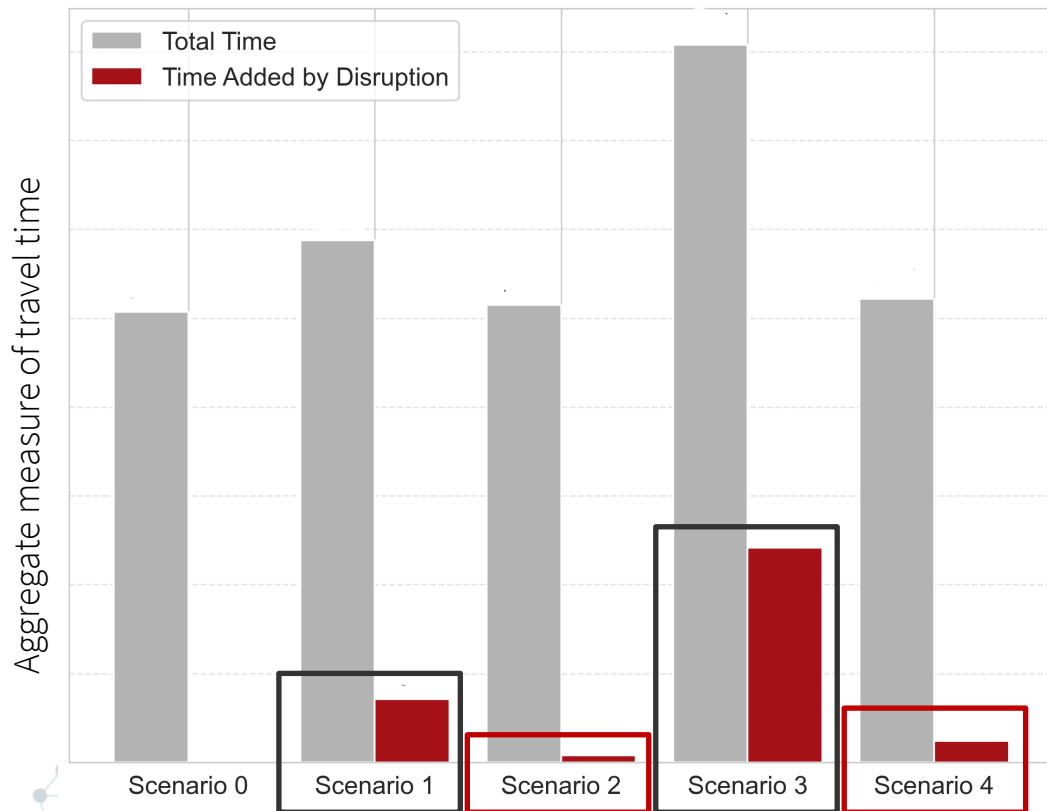
Disinfo + transit: illustrative example

We identified the top 10% most central nodes to form the subset of stations that disinformed users avoid

We conducted a sensitivity analysis on the threshold parameter that determines whether a user is classified as disinformed

Scenario	Threshold	Protection allowed
0	<i>No disruption</i>	
1	0.5	No
2	0.5	Yes
3	0.3	No
4	0.3	Yes

Disinfo + transit: illustrative example

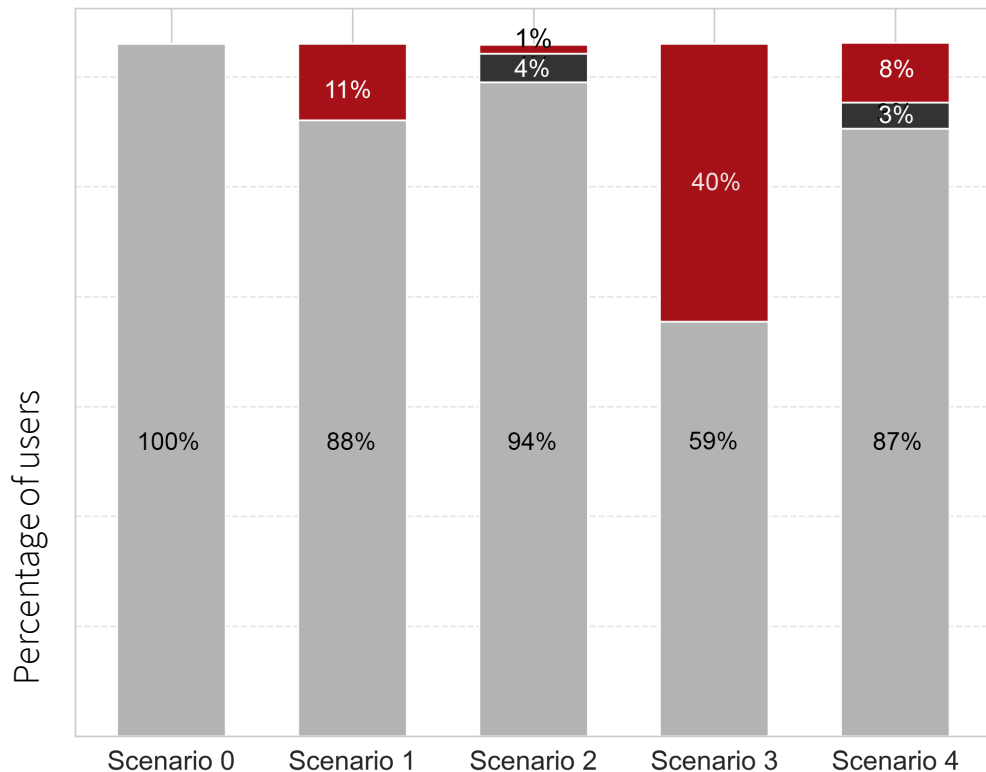


Reducing the threshold required to believe disinformation results in a 340% increase in added travel time in the absence of a protection strategy

Protection helps substantially regardless of threshold

Disinfo + transit: illustrative example

■ Informed Users (Unprotected) ■ Protected Users ■ Disinformed Users



Even with a relatively small number of protected users, the reduction in the effects of disinformation is substantial

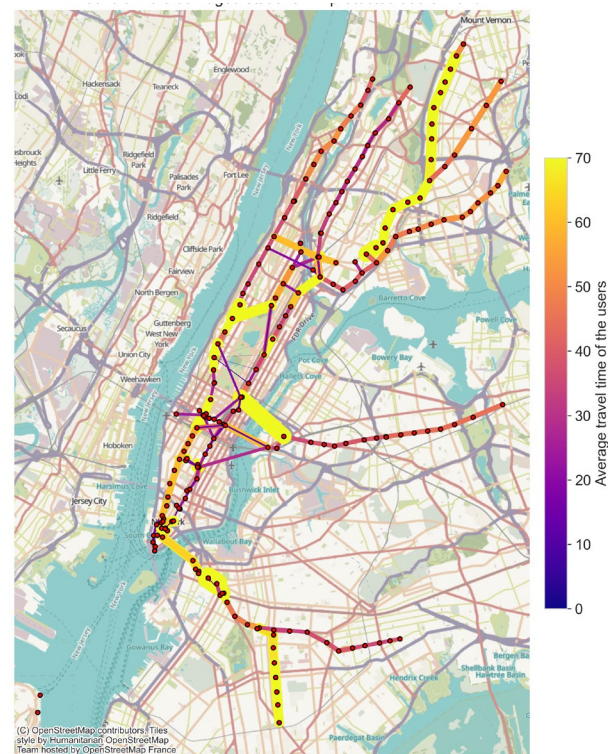
Disinfo + transit: illustrative example



Scenario 0



Scenario 1
threshold 0.5, no protection



Scenario 2
threshold 0.5, with protection

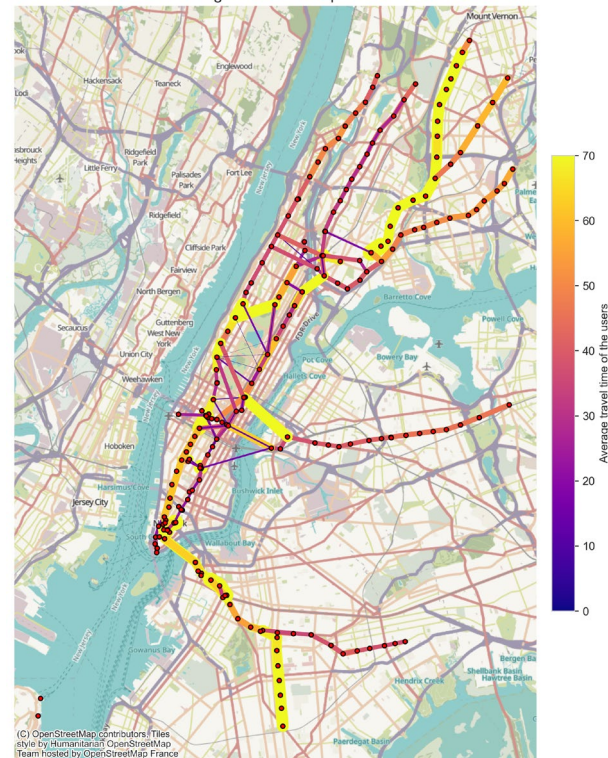
Disinfo + transit: illustrative example



Scenario 0



Scenario 3
threshold 0.3, no protection



Scenario 4
threshold 0.3, with protection

Weaponized disinformation + transit

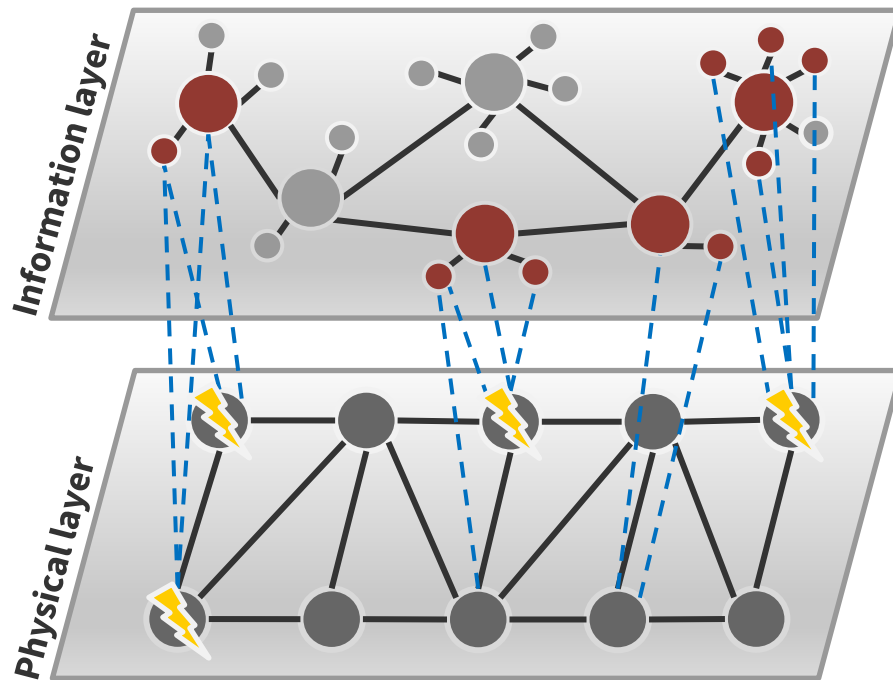
Future ideas

Protecting both layers

Tradeoffs between protecting the physical layer versus curbing the spread of disinformation

Multi-level models

More effectively representing “attacker-defender” relationships



Disinformation, future grid

Rapid electric grid decarbonization and 2-3 times the current solar and wind energy capacity deployment are key to keeping the global temperature rise below 2°C

e.g., the Biden administration had targeted a fossil fuel free electricity sector by 2035

Disinformation, future grid

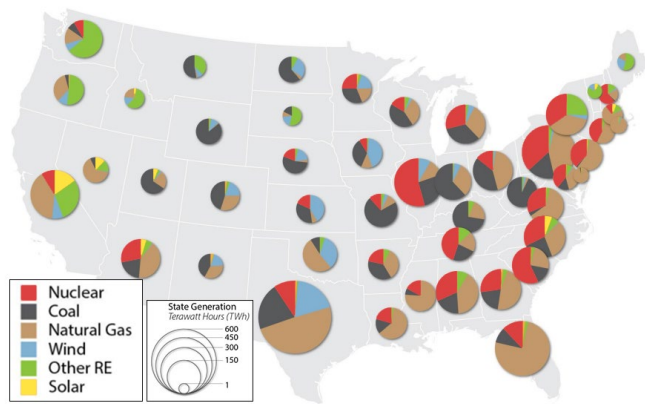
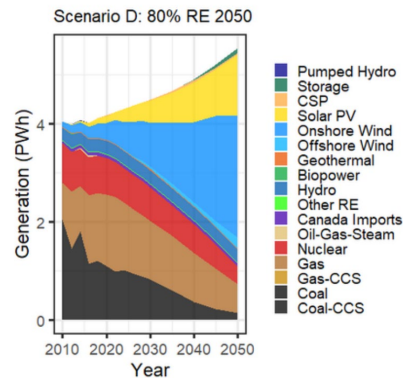
However, misinformation can alter public perceptions of the benefits of cleaner sources of energy, and could potentially undermine the energy transition

The US Department of Energy *called* such misinformation “a key threat to decarbonizing the grid” because “misinformation is raising doubts about renewable energy and slowing or derailing projects”

Disinformation, future grid

How does
mis/disinformation impact
future energy portfolios?

And how does that impact
different vulnerable
communities?



Background

Recent work

Our activities

Concluding remarks



Weaponized disinformation

Special issue of *Risk Analysis*!

Risk Implications of Mis/Disinformation

Submission deadline: July 1, 2025 (or so)

My thanks to...

- The Aalto Science Institute
- The Aalto University Department of Mathematics and Systems Analysis
- The Systems Analysis Laboratory
- Ahti Salo
- Jean Sibelius, Karl Fazer

End of presentation



kashbarker@ou.edu



www.ou.edu/systemslab



@kashbarker



RISK-BASED
SYSTEMS ANALYTICS
LABORATORY



CENTER FOR
CYBER•PHYSICAL•SOCIAL
SYSTEMS