

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 29-08-2021		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 14-Jun-2019 - 15-Mar-2021	
4. TITLE AND SUBTITLE Final Report: Resource allocation and statistical estimation in epidemic networks: Scalable algorithms and analysis			5a. CONTRACT NUMBER W911NF-19-1-0343		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHORS			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Wisconsin - Madison Suite 6401 21 N Park Street Madison, WI 53715 -1218			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 73724-CS-YIP.2		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Po-Ling Loh
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 608-262-5505

RPPR Final Report

as of 19-Oct-2021

Agency Code: 21XD

Proposal Number: 73724CSYIP
INVESTIGATOR(S):

Agreement Number: W911NF-19-1-0343

Name: Po-Ling J. Loh
Email: loh@ece.wisc.edu
Phone Number: 6082625505
Principal: Y

Organization: **University of Wisconsin - Madison**

Address: Suite 6401, Madison, WI 537151218

Country: USA

DUNS Number: 161202122

EIN: 396006492

Report Date: 15-Jun-2021

Date Received: 29-Aug-2021

Final Report for Period Beginning 14-Jun-2019 and Ending 15-Mar-2021

Title: Resource allocation and statistical estimation in epidemic networks: Scalable algorithms and analysis

Begin Performance Period: 14-Jun-2019

End Performance Period: 15-Mar-2021

Report Term: 0-Other

Submitted By: Po-Ling Loh

Email: loh@ece.wisc.edu

Phone: (608) 262-5505

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees:

STEM Participants: 2

Major Goals: Our proposed research will advance the frontiers of stochastic modeling for dynamic networks. We will analyze problems in influence maximization, fractional immunization and boosting, and detection and estimation, with the unifying theme of devising computationally efficient algorithms with rigorous mathematical guarantees for allocating resources and estimating characteristics of time-varying networks. This work significantly generalizes and extends existing approaches in network science by allowing greater flexibility in modeling epidemics and interventions. In particular, we will depart from standard submodularity assumptions known to be irreflexive of real-world cascades, and allow targeted interventions to be fractional rather than binary. We will also develop algorithms suited to incorporate uncertainty about network connectivity, and provide methods for estimating graph characteristics based on sensor data.

Accomplishments: Two directions were primarily explored in this project. The first was most closely aligned with the original plan for the project. One PhD student worked on developing new methods for resource allocation in networks under a "fractional immunization" framework. This mostly consisted of formulating the allocation problem as a constrained optimization problem in terms of eigenvalues of the underlying graph structure. The student then studied the performance of the proposed algorithm under various network structures, focusing on optimality and stability of the algorithm. Analogies in the "fractional boosting" problem were also explored. However, work on this project ended up stalling for quite a while, due to personal and family problems that the student encountered after the pandemic began.

The second student who was recruited to help with the project had prior interest in the theory of deep learning. Thus, he worked on a problem which was more abstractly connected to the goals of the project. It is unclear whether existing models of statistical robustness would be appropriate to capture robustness of discrete structures such as networks. The student studied the mathematical framework of Wasserstein contamination. As it turned out, neural networks were instrumental in designing computationally friendly estimators which were robust to this type of contamination. After finishing the theoretical work, the student began exploring how neural networks might be used for resource allocation in graphs.

Training Opportunities: The work in this project contributed directly to the training of two PhD students, Michael Liou and Zheng Liu, who worked on the fractional immunization and neural network projects, respectively.

RPPR Final Report

as of 19-Oct-2021

Results Dissemination: In 2021, the PI gave several invited talks on her work on networks, including the Joint Statistical Meetings (JSM) conference; the American Mathematical Society (AMS) sectional meeting; and the Joint Probability Seminar at the Universities of Frankfurt/Mainz/Darmstadt.

The paper on robustness of neural networks has been posted to the arXiv.

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: Graduate Student (research assistant)

Participant: Michael Liou

Person Months Worked: 5.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Zheng Liu

Person Months Worked: 12.00

Funding Support:

Project Contribution:

National Academy Member: N

ARTICLES:

Publication Type: Journal Article

Peer Reviewed:

Publication Status: 4-Under Review

Journal: Information and Inference

Publication Identifier Type:

Publication Identifier:

Volume:

Issue:

First Page #:

Date Submitted:

Date Published:

Publication Location:

Article Title: Robust W-GAN-Based Estimation Under Wasserstein Contamination

Authors: Zheng Liu, Po-Ling Loh

Keywords: robustness, deep learning, minimax optimality

Abstract: Robust estimation is an important problem in statistics which aims at providing a reasonable estimator when the data-generating distribution lies within an appropriately defined ball around an uncontaminated distribution. Although minimax rates of estimation have been established in recent years, many existing robust estimators with provably optimal convergence rates are also computationally intractable. In this paper, we study several estimation problems under a Wasserstein contamination model and present computationally tractable estimators motivated by generative adversarial networks (GANs). Specifically, we analyze properties of Wasserstein GAN-based estimators for location estimation, covariance matrix estimation, and linear regression and show that our proposed estimators are minimax optimal in many scenarios. Finally, we present numerical results which demonstrate the effectiveness of our estimators.

Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: Y

RPPR Final Report
as of 19-Oct-2021

Partners

,

I certify that the information in the report is complete and accurate:

Signature: Po-Ling Loh

Signature Date: 8/29/21 7:19AM

Convex Optimization Framework for Influence Maximization and Fractional Immunization over a Network

Michael Liou

University of Wisconsin-Madison, Department of Statistics, Madison, WI, USA

E-mail: liou@stat.wisc.edu

Po-Ling Loh

University of Wisconsin-Madison, Department of Statistics, Madison, WI, USA

E-mail: ploh@stat.wisc.edu

Summary. Abstract goes here

Keywords: epidemiology, convex optimization, influence maximization

1. Introduction and Motivation

Recently there has been great interest in modeling dynamic spread of a contagion across either a technological or social network. We are motivated by two scenarios, one in which we wish to contain the spread of a virus in a time-sensitive manner. The other scenario in which we wish for information to be distributed as quickly and as widely as possible across a network. Modeling these dynamics over a probabilistic network data structure is a strong choice in that we are able to capture information of differing probabilities of transmission between different units. We are also motivated by the potential of making interventions in both situations to help achieve the desired level of spread. In both scenarios, we are operating with a limited amount of resources to distribute to the units, with the idea that these resources are able to increase or decrease the transmission rates through that particular units on the network, we can

think of these resources as either medical supplies, campaign personnel, funds. We will assume that we have some prior knowledge of about the likelihood of transmission with and without interventions. If we are modeling this network in a geospatial context, this prior knowledge could be some function of distance, or how well connected two locations are (how many airline flights travel between these major cities). Furthermore, since the event horizon for such dynamics over a network may be uncertain, we consider how to intelligently reallocate these resources at discrete time points as well.

In trying to minimize or maximize the transmission across a network by allocating resources, we can formulate this as an optimization problem. These problems have been previously studied under largely separate domains, commonly known as problems in control theory as “influence maximization” or “immunization in networks”. In the literature of immunization of networks, this context of problem has been studied with slight modifications of whether directionality of transmission, and under various epidemic Susceptible-Infected related stochastic models. Focus in the influence maximization has been on complexity theory or computational algorithms. There has been little work at the intersection of these two fields, yet, there is good reason to believe there is a strong relationship between these two dynamics across a network. Should the same units on a network deemed important for targeting in both objectives? Does the amount of resources that we spend at each of the nodes may be more dependent on the structure of the network rather than the rates of transmission? These are just a few of the many questions we hope to answer, but we must have some way of relating these two objectives in order to adequately study these interactions. Here we offer a unified convex framework to relate these dynamics in terms of optimal resource allocation.

The organization of the following will be a formulation of the problem with notation used in the next section. Next, we examine some of the key results that our framework utilizes. Finally, we discuss the simulation results and the implications for controlling network dynamics.

2. Notation and Preliminaries

2.0.1. Graph Theory

Let the weighted edge, undirected, no self loops graph, \mathcal{G} be denoted by the triple $(\mathcal{V}, \mathcal{E}, \mathcal{W})$ defined to be the set of vertices, edges, and edge weights respectively. We have $\mathcal{V} \triangleq \{v_1, v_2, \dots, v_n\}$ with $|V| = n$ nodes and $|E| = m$ edges. The related adjacency matrix of a graph G is denoted by the $A_G \triangleq [a_{ij}]_{n \times n}$. In which each entry a_{ij} in the matrix is the weight of the edge between the nodes v_i and v_j . We then denote the eigenvalues of a square (symmetric) matrix A by $\lambda_1(A), \lambda_2(A), \dots, \lambda_n(A)$ in decreasing order. At first, we consider the structure of the graph to be fixed throughout the time period we are considering, and we will discuss relaxing this assumption later.

[[Perhaps more details of eigenvalue properties we use in the proofs to come.]]

2.0.2. Epidemiology Models

There are a couple of epidemiological models that we consider for the dynamics of minimizing the spread over a network. Among the most popular are the compartmental models, in which each unit belongs to some category, either a susceptible group (those that still can get the virus), or an infected group (those that have been infected by the virus and can transmit to others. This abbreviated is known as an SI model. We could further extend this with another category known as the “recovered” group (those that have recovered from the virus, are unable to get it again, and also will not transmit). This is known as the SIR model.

We study the network dynamics with a mean-filed approximation, which is commonly used in this field due to the fact that the number of states that this stochastic process can take grows exponentially with the number of nodes in the network.

[[Perhaps more details needed of established results in these fields, for example the markov differential equation, and stochastic approximation]]

We only consider the possibility of a node to be affected or unaffected, there is no partial state for each unit in the node, however each node may be allowed to receive a fractional amount of intervention. In fractional immunization, we intervene at a certain time t , and have some budget C_t to allocate to each of the nodes, with some overall budget $C = \sum_t C_t$ over

the entire time horizon. In real scenarios, the distribution of resources to one area of the network may be more difficult than distributing resources to another area. Furthermore, there may be a scaled effort in that once some resources are distributed to one unit of the network, it costs less to provide resources to a node that has already received resources compared to allocating the same amount of resources to a new unit in the network. To account for the flexibility of these scenarios, we allow for each unit in the network to have a monotonically increasing/decreasing function (depending on the problem objective) $f_i(\beta_i)$ describing how efficiently we can distribute resources to this area and how much impact those resources would have in changing the transmission rate. Furthermore, each node will also allow for the possibility of a distinct lower and upper bound on the amount that an intervention would help. That is, these bounds signify what the transmission rate look like for that unit in the network if it had an infinite amount of resources.

The convex optimization framework depends on an exponential stability condition, which serves as a restriction on the large eigenvalue of a matrix relating the transmission rates and the network structure. In the context of fractional immunization, we carry forward a similar framework as outlined by Preciado et al. (2014).

$$\begin{aligned}
 & \text{maximize} && \varepsilon \\
 & \text{subject to} && \lambda_1(BA_G) < -\varepsilon \\
 & && \underline{\beta}_i \leq \beta_i \leq \bar{\beta}_i \\
 & && \sum_i f_i(\beta_i) < C
 \end{aligned}$$

This optimization problem is extended with slight modification to the SIR model under the following optimization problem, where δ_i denotes the recovery rates of each node v_i , and $g_i(\delta_i)$ denotes the monotonically increasing efficiency function of distributing resources to node v_i .

$$\begin{aligned}
 & \text{maximize} && \varepsilon \\
 & \text{subject to} && \lambda_1(\text{diag}(\beta_i)A_G - \text{diag}(\delta_i)) < -\varepsilon \\
 & && \underline{\beta}_i \leq \beta_i \leq \bar{\beta}_i \\
 & && \underline{\delta}_i \leq \delta_i \leq \bar{\delta}_i \\
 & && \sum_i f_i(\beta_i) + g_i(\delta_i) < C
 \end{aligned}$$

The formulation of the opposite problem of maximal influence in this convex optimization statement.

$$\begin{aligned}
 & \text{minimize} && \lambda_1(\text{diag}(\beta_i)A_G) \\
 & \text{subject to} && \underline{\beta}_i \leq \beta_i \leq \bar{\beta}_i \\
 & && \sum_i f_i(\beta_i) < C
 \end{aligned}$$

[Proof that both of these frameworks are convex. In short, the second one because we're simply minimizing the spectral radius of a symmetric matrix, proof by perron-frobenius, also in convex optimization book. Proof of the other is somewhat similar.]

[TODO: Discuss the implications of time in this framework.] [TODO: Discuss the overlap of nodes targeted]

3. Results

3.1. Simulation Results

Fig. 1. Graphic showing intersection of identified nodes of importance for both objectives

Fig. 2. Graphic showing infection across many network topologies

3.2. *Stability*

In this section, we discuss some of the stability problems of such an optimization procedure. We note specifically that the solution to both optimization problems is unstable when there is not enough resources to be distributed. In such a scenario, the takeaway is that no matter how the resources are allocated, there will be very little impact on the overall dynamics of the problem if any at all.

3.3. *Real Dataset Results*

4. **Conclusions**

5. **Appendix**

5.1. *Proofs*

References

Preciado, V. M., Zargham, M., Enyioha, C., Jadbabaie, A. and Pappas, G. J. (2014) Optimal resource allocation for network protection against spreading processes. *IEEE Transactions on Control of Network Systems*.