Look-Ahead Robust Network Optimization with Generative State Predictions

Fei Xu Yu[1], Zuyuan Zhang[1], Emily Grob[1], Gina Adam[1], Sean Coffey[2], Nathaniel D. Bastian[2], Tian Lan[1]

[1]George Washington University, Washington, DC USA
[2]United States Military Academy, West Point, NY USA
{fxyu, zuyuan.zhang, emily.grob, ginaadam, tlan}@gwu.edu, {sean.coffey, nathaniel.bastian}@westpoint.edu

Abstract

In network optimization, especially in environments with significant uncertainty, traditional approaches are often overly conservative which leads to inefficiencies. Robust network planning is essential when network conditions can vary unpredictably. Traditional robust approaches tend to protect against all types of uncertainty, including those with minimal impact, resulting in suboptimal performance and resource allocation. To address this, we propose leveraging diffusion models to generate future network states based on historical data, capturing realistic variations without overgeneralizing uncertainty. Our method defines an uncertainty set based on these generated states, focusing on probable scenarios rather than extreme outliers. Using this uncertainty set, we use robust optimization to allocate resources and ensure network reliability under dynamic conditions. Preliminary experiments demonstrate that our approach achieves a balance between robustness and efficiency, significantly outperforming rational methods in realistic network scenarios.

Introduction

Optimization theory has been applied to solve a wide range of problems in communication and networking (Yang et al. 2008; Sedghi, Ahmadian, and Aliakbar-Golkar 2016), from resource allocation (Chen, Ling, and Giannakis 2017; Halabian 2019) to robust power control (Shen, Dai, and Win 2014). Traditional optimization assumes static, precise data, but real-world networking often involves uncertain conditions, leading to suboptimal or infeasible solutions in practice.

Robust optimization has gained attention as a methodology specifically to address data uncertainty in network optimization problems. Robust optimization (EI-Ghaoui and Lebret 1997; Ben-Tal and Nemirovski 1999; Nemirovski 2003) seeks solutions that maintain feasible and near-optimal solutions across a range of perturbations in the nominal problem's parameters. Each robust optimization problem is typically defined by three key elements: a nominal formulation, a robustness criterion, and an uncertainty set. This approach transforms a nominal optimization problem into a robust one, while preserving the essential properties, such as convexity. Robust network planning is vital in contested environments (Szabo et al. 2020), where network conditions can vary unpredictably. However, traditional robust approaches tend to protect against all types of uncertainty, which leads to either inefficient or excessive protection.

In this work, we propose a novel approach that leverages diffusion models (Ho, Jain, and Abbeel 2020) to address data uncertainty through synthetic data generation. By training diffusion models on historical network data, we predict future network conditions, such as variations in the signal-to-interference-plus-noise ratio (SINR), terrain changes, environmental factors, and adversarial behaviors, to define a 95% confidence interval for a representative uncertainty set. Using this set, We apply a robust optimization formulation to develop robust and effective solutions for network optimization like bandwidth allocation and rate control.

Our experiment in a simulated network environment shows that our approach, focused on maximizing the minimum data rate for each transmitter and receiver, significantly outperforms the existing methods, like Yang et al. (2008), in terms of both efficiency and robustness.

Related Work

Existing approaches of robust network optimization techniques such as KL divergence (Hu and Hong 2013) or ellipsoid sets (Yang et al. 2008) are too general and do not capture the specific patterns seen in real-world scenarios. Some papers use a max-min formulation (Tu, Chen, and Yue 2024). As a result, existing algorithms tend to be overly pessimistic by protecting against all uncertainties. Machine learning has also been applied in network planning (Wang et al. 2018) to predict and classify network traffic (Zhang et al. 2014; Chen, Wen, and Geng 2016), and allocate resources (Winstein and Balakrishnan 2013; Mao et al. 2017). Despite their strength, these methods often have limited interpretability and robustness when dealing with unseen events or uncertainty in the problem. Recently, diffusion models (Ho, Jain, and Abbeel 2020) have demonstrated significant potential in generative tasks, particularly for modeling complex data distributions. In planning, diffusion models (Ajay et al. 2023; Janner et al. 2022) have been used to generate future states conditioned on past data. Our work integrates these three areas by proposing a novel framework, generative robust optimization, to solve network optimization problems. Unlike traditional approaches that rely on static or conservative uncertainty sets, our method generates a scenario-specific uncertainty set through diffusion models.

Methodology

Problem Statement

In this work, we address a general robust optimization problem aimed at optimizing network resources in dynamic and uncertain environments. Specifically, we focus on scenarios where future network states are critical to achieving the optimal resource allocation. The main challenge in these scenarios is the inherent uncertainty of network conditions, which can fluctuate due to factors such as adversarial interference, environmental changes, and user mobility. Robust optimization is essential in these settings to ensure that the solution remains effective across a range of possible future states.

Many robust network optimization problems are often formulated as an optimization with a concave objective function f_0 (e.g., data rate and power consumption) over a set of linear constraints (e.g., SINR and bandwidth constraints):

$$\begin{array}{ll} \text{maximize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \forall (a_i, b_i) \in A \\ \text{variables} & x \end{array} \tag{1}$$

where (a_i, b_i) lies in a compact uncertainty set A (e.g., corresponding to possible network operating conditions). This formulation captures many important network design problems, such as maximizing minimum data rate under bandwidth constraints and minimizing power consumption under SINR constraints (Yang et al. 2008). For instance, consider a network with a set L of communication links. Let $\text{SINR}_{(p,h)}^{\ell}$ be the SINR function of link ℓ depending on the transmit powers p and the channel gains/conditions h. For an uncertainty set H of channel gains, a robust bandwidth allocation problem aims to maximize the minimum sum rate over different $h \in H$, i.e.,

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & R \leq r_h, \ \forall h \in H \\ & r_h = \displaystyle\sum_{\ell \in L} b_\ell \log(1 + \text{SINR}_{(p,h)}^{\ell}), \ \forall h \in H \\ & \displaystyle\sum_{\ell} b_\ell \leq B. \\ \text{Variables} & b_{\ell h}, r_h \geq 0, \ \forall \ell, h \end{array} \tag{2}$$

where B is the total bandwidth constraint, b_ℓ is the bandwidth assigned to link ℓ, and r_h is the sum data rate achieved achieved in scenario $h \in H$ from the uncertainty set H.

A key assumption in robust optimization is the availability of an uncertainty set that can capture the range of possible states the network may encounter. However, in real-world applications, future states are unknown and difficult to predict with precision. The lack of future conditions presents a challenge, as standard robust optimization approaches rely on predefined uncertainty sets. To address this, we introduce a data-driven approach using diffusion models to generate realistic samples of potential future states. Diffusion models, which are generative models trained to learn the distribution of high-dimensional data, provide a powerful tool to model and generate complex scenarios that reflect real-world variability in network conditions.

Our approach proceeds in two states. First, we use a diffusion model to generate future network states based on historical data, capturing variations in SINR. This generated data serves as the uncertainty set, providing a basis for robust optimization that reflects realistic and scenario-specific variations. Second, we apply a robust optimization formulation that leverages this uncertainty set to ensure that our solution is resilient across the majority of possible future states. To avoid an overly conservative uncertainty set, we focus on the 95% confidence range of the uncertainty set. This will allow us to find feasible and near-optimal solutions while maintaining a balance between flexibility and reliability.

Generation with Diffusion Models

In generation using diffusion models (Ho, Jain, and Abbeel 2020), $q(\tau_0)$ is the data distribution and $p(\tau_N)$ is a Gaussian prior. Plan generation is formulated as trajectory generation τ through a learned iterative denoising diffusion process $p_\theta(\tau_{i-1}|\tau_i) := \mathcal{N}(\tau_{i-1}; \mu_\theta(\tau_i, i), \Sigma_\theta(\tau_i, i))$. This denoising process reverses a forward diffusion process $q(\tau_i|\tau_{i-1}) := \mathcal{N}(\tau_i; \sqrt{(1-\beta_i)\tau_{i-1}}, \beta_i I)$ that adds noise to trajectories, where $\beta_1, \ldots, \beta_N$ are predetermined noise coefficients. The N-step diffusion process in both directions is modeled as Markov transition probabilities:

$$\begin{aligned} q(\tau_{0:N}) &:= q(\tau_0) \prod_{i=1}^{N} q(\tau_i|\tau_{i-1}), \\ p_\theta(\tau_{0:N}) &:= p(\tau_N) \prod_{i=1}^{N} p_\theta(\tau_{i-1}|\tau_i). \end{aligned} \tag{3}$$

Diffuser maximizes trajectory likelihood through a variational lower bound:

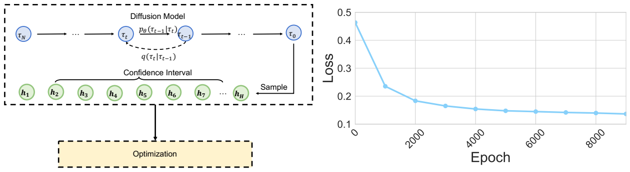
$$\begin{aligned} \mathbb{E}_{q(\tau_0)}[\log p_\theta(\tau_0)] &\geq \mathbb{E}_{q(\tau_{0:N})} \left[\log \frac{p_\theta(\tau_{0:N})}{q(\tau_{1:N}|\tau_0)} \right] \\ &\approx \mathbb{E}_{\tau_{1:N} \sim q(\tau_{1:N})} \left[\sum_{i=1}^{N} \log p_\theta(\tau_{i-1}|\tau_i) \right]. \end{aligned}$$

Trajectories are generated by sampling iteratively from the denoising process for N steps. Following Decision Diffuser (Ajay et al. 2023) to perform conditional diffusion, we learn and sample from the conditional trajectory density $p_\theta(\tau|J)$.

Our Generative Robust Optimization

With the assumption that the uncertainty set A is compact, we can introduce an auxiliary function $g_i(x) = \max_{a_i \in A}(a_i - \overline{a})^T x$ and rewrite the optimization in (1) as

$$\begin{array}{ll} \text{maximize} & f_0(x) \\ \text{subject to} & \overline{a}_i^T x + g_i(x) \leq b_i, \forall (a_i, b_i) \in A \\ \text{variables} & x \end{array} \tag{4}$$



(a) (b)

Figure 1: (a) shows the pipeline of our generative robust optimization. (1) We use the simulated channel gains as input to the diffusion model. (2) We train the diffusion model to generate future channel gains based on history. (3) We apply the predicted channel gains to optimization to solve network problems. (b) shows the convergence of the diffusion model.

where $\bar{a}$ is a nominal value, e.g., the mean. As shown in Yang et al. (2008), this leads to a fully decentralized solution to robust optimization. In practice, obtaining the uncertainty set A is often challenging due to the lack of accurate knowledge about the underlying distribution of parameters. Existing approaches typically rely on protection regions to approximate the uncertainty. However, these methods may fail to capture the true dynamics and variability of real-world scenarios. We will leverage the problem formulation in (4) and use the proposed diffusion model to generate the uncertainty set A. Our solution overcomes this limitation by using a diffusion model to approximate A_i as a sample from a learned distribution $q(\tau_0)$, as shown in figure 1a. Our approach ensures capturing realistic variations in the environment, providing a more flexible and effective solution to robust optimization.

Specifically, we generate an uncertainty set of possible channel gains from the diffusion model, i.e., $h \sim q(\tau_0)$. We then formulate the sum data rate as a function of power p and channel gain h using $r_h = \sum_{\ell \in L} b_\ell \log(1 + \mathrm{SINR}^{\ell}_{(p,h)})$ and rewrite the robust network optimization in (2) as follows:

$$\begin{aligned} \text{maximize} \quad & R && (5) \\ \text{subject to} \quad & R \le r_h^{norm} + g_h(b_1, \ldots, b_L),\ \forall h \\ & g_h(b_1, \ldots, b_L) = \min_{h \sim [q(\tau_0)]} (r_h - r_h^{norm}),\ \forall h \\ & \sum_{\ell} b_\ell \le B \\ \text{Variables} \quad & b_{ij}, r_i \ge 0 \quad \forall i, j \end{aligned}$$

where r_h^{norm} is a nominal mean data rate in scenario h. Further, we sample h from the 95% confidence interval of the diffusion model's output distribution, i.e., $h \sim [q(\tau_0)]_{95\%}$. This means that extreme outliers are excluded from the uncertainty set, ensuring that the optimization is not overly conservative. It also achieves different levels of robustness in planning. We leverage the distributed algorithm in (Yang et al. 2008) to solve this robust optimization. We omit the details due to space limitations.

Experiment

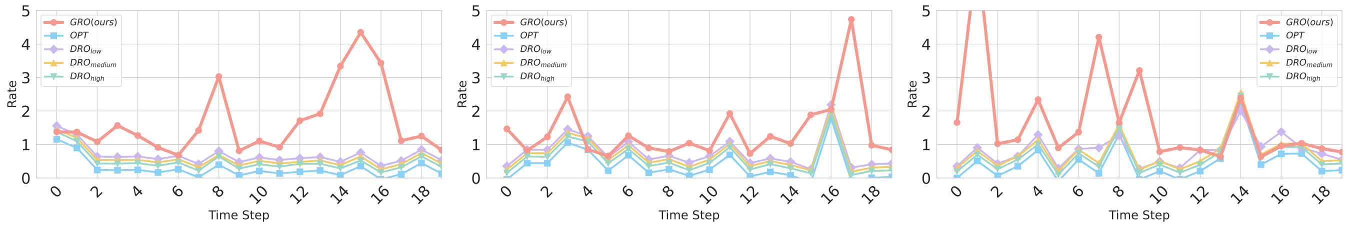
In this section, we evaluate the performance of our proposed generative robust optimization (GRO) framework. We first show how to obtain the channel gain h through diffusion model. Then, we compare it to traditional robust optimization methods that use a single time step to sample data. We analyze how the minimum data rate R changes with different protection regions, highlighting the benefits of our diffusion-based uncertainty modeling. Next, we explore the trade-off between robustness and performance by testing the GRO framework with varying confidence intervals. This analysis demonstrates the impact of choice on the confidence level. It provides insights into the balance between robustness and conservatism. All experiments are conducted on a Linux machine with AMD EPYC 7513 32-Core Processor CPU and an NVIDIA RTX A6000 GPU, implemented in python3 and compiled by a Python compiler.

Diffusion Model Training The input to the diffusion model is generated by a simple state-based network simulator acting as the channel model. The channel model simulates the movement of nodes within a dynamic network under configurable conditions. The channel is represented as a grid of randomly placed nodes, categorized into three types: soldiers, drones, and adversaries. The channel model includes configurable parameters such as the number of nodes of each type, the probability of node movement, the transmission power, the speed of nodes, and the height range for drones. Movement patterns for each node are modeled as independent Bernoulli processes, with each node having a configurable probability of movement per time step. Channel gains are computed based on the proximity of adversary nodes and soldier nodes to drone nodes, i.e., Gain $= \frac{1}{d^\alpha}$, where d is the distance between nodes and α represents the path loss exponent. For drones, α is set to 2, while for soldiers and adversaries, α is set to 3. The simulation is designed to emulate scenarios where soldiers and drones need to maintain a communication link in dynamic and contested environments, where the adversaries attempt to degrade communication by introducing interference. The drones must dynamically adapt bandwidth allocation to maximize the minimum data rate for soldiers.

To simulate dynamic network conditions, we simulate 100 time steps of the network states. The data are then used as input to the diffusion model. It will be trained to learn the underlying patterns in the data and generate realistic future distance matrices based on history. Figure 1b illustrates the training loss curve of the diffusion model.

Comparison with Baselines

We compared our proposed method of generative robust optimization with standard optimization and distributed robust optimization with three different protection regions. For standard optimization OPT, we do not consider uncertainty set in Equation (2). For distributed robust optimization, we construct uncertainty sets by varying the data in an interval of [data-x%, data+x%]. We chose $x = 5\%$, $x = 10\%$, and $x = 20\%$ for DRO$_{low}$, DRO$_{medium}$, DRO$_{high}$ respectively. We sampled 20 channel gains h from the interval for each soldier-drone pair and computed the corresponding SINR values. Using these values, we applied distributed robust optimization to optimize the minimum data rate R. For the outage probability $p(r < z)$, we choose the threshold to be $z =$ 5th percentile of the data rates in GRO to showcase our



(a) Run 1 (b) Run 2 (c) Run 3

Figure 2: Comparison of our GRO with baselines in 3 different runs. Our method is able to improve the minimum data rate and demonstrates robust performance across different scenarios. The outage probability is significantly reduced.

method's ability to protect against the worst cases.

As shown in Table 1, our generative robust optimization approach improves the mean and variance of the data rate and our method achieves a better balance between robustness and performance. We also see a significant decrease in the outage probability of data rates that fall below the threshold z. This demonstrates the effectiveness of generative robust optimization in managing uncertainty while ensuring higher minimum data rates.

In Figure 2, we present the results of three experimental runs conducted under different network scenarios. The results demonstrate that our generative robust optimization consistently outperforms the alternative approaches across the majority of the time steps, highlighting its effectiveness and robustness in varying conditions.

Method	Mean μ_r	Variance σ_r^2	Outage $p(r < z)$
OPT	1.1892	30.9422	27.5000%
DRO$_{low}$	1.2864	30.9574	27.5000%
DRO$_{medium}$	1.2815	30.8134	27.500%
DRO$_{high}$	1.2838	30.9662	27.5000%
GRO (ours)	**1.4636**	**20.5275**	**5.000%**

Table 1: This illustrates the average, variance, and probability of the data rate r_i is below a threshold for each soldier i under different optimization methods.

CI	Mean μ_r	Variance σ_r^2	Outage $p(r < z)$
95%	1.1889	22.7073	5.0000%
90%	1.1889	22.7073	5.0000%
80%	1.9449	94.8504	5.0000%
50%	2.0629	97.0007	7.5000%

Table 2: This illustrates the average, variance, and probability of the data rate r_i is below a threshold for each soldier i under different confidence intervals (CI). It shows the trade-off between the mean and variance of data rates.

Trade-off Analysis. We further analyze the trade-off between the confidence intervals (CI) using our generative robust optimization approach. The results are shown in Table 2. By varying the confidence intervals, we evaluate how the level of uncertainty affects the data rate. Under high confidence intervals (95% and 90%), the uncertainty set includes a wide range of potential scenarios, capturing both typical variations and moderate outliers. This ensures the optimization accounts for the most realistic conditions, resulting in robust data rate solutions across the network. The similarity between 95% and 90% indicates that the additional scenarios captured by the extra 5% have little impact on the optimization. These rare cases have minimal influence on the optimization process. With an 80% confidence interval, the uncertainty set starts to exclude some realistic variations. The optimization focuses on a narrower set of scenarios, which reduces robustness. The average data rate increases slightly but the variance increases significantly. With a 50% confidence interval, the uncertainty set captures only the central portion of the data distribution, excluding many real-world variations. This leads to overly optimistic solutions that fail to account for common uncertainties, leading to a further increase in the average and the variance of data rates. This trade-off leads to more data rates falling under the threshold z. This analysis demonstrates that 90%-95% confidence intervals provide the best result by capturing the majority of realistic conditions without rare outliers.

Conclusion

The development of robust optimization models for communication network design is important in addressing uncertainty in real-world scenarios. In this paper, we introduced a novel approach that integrates diffusion models with robust optimization to tackle the challenges posed by dynamic environments. We applied our approach to a bandwidth allocation problem involving soldiers and drones, aiming to maintain a minimum data rate under uncertain SINR conditions. Using diffusion model to generate scenario-specific uncertainty sets, our GRO methods demonstrated superior performance and robustness compared to traditional methods. Additionally, we explored the trade-off between different confidence intervals. The results highlight our approach's ability to balance robustness and performance, making it suitable for many communication network applications. This study lays the groundwork for further innovations in network optimization that bridge machine learning and robust optimization.

Acknoledgement

This work was supported in part by the U.S. Military Academy under Cooperative Agreement No.W911NF-22-2-0089. The views and conclusions expressed in this paper are those of the authors and do not reflect the official policy or position of the U.S. Military Academy, U.S. Army, U.S. Department of Defense, or U.S. Government.

References

Ajay, A.; Du, Y.; Gupta, A.; Tenenbaum, J.; Jaakkola, T.; and Agrawal, P. 2023. Is Conditional Generative Modeling all you need for Decision-Making? arXiv:2211.15657.

Ben-Tal, A.; and Nemirovski, A. 1999. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1): 1–13.

Chen, T.; Ling, Q.; and Giannakis, G. B. 2017. An Online Convex Optimization Approach to Proactive Network Resource Allocation. *IEEE Transactions on Signal Processing*, 65(24): 6350–6364.

Chen, Z.; Wen, J.; and Geng, Y. 2016. Predicting future traffic using Hidden Markov Models. In *2016 IEEE 24th International Conference on Network Protocols (ICNP)*, 1–6.

EI-Ghaoui, L.; and Lebret, H. 1997. Robust solutions to least-square problems to uncertain data matrices. *Sima Journal on Matrix Analysis and Applications*, 18: 1035–1064.

Halabian, H. 2019. Distributed Resource Allocation Optimization in 5G Virtualized Networks. *IEEE Journal on Selected Areas in Communications*, 37(3): 627–642.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33: 6840–6851.

Hu, Z.; and Hong, L. J. 2013. Kullback-Leibler divergence constrained distributionally robust optimization. *Available at Optimization Online*, 1(2): 9.

Janner, M.; Du, Y.; Tenenbaum, J. B.; and Levine, S. 2022. Planning with Diffusion for Flexible Behavior Synthesis. arXiv:2205.09991.

Mao, B.; Fadlullah, Z. M.; Tang, F.; Kato, N.; Akashi, O.; Inoue, T.; and Mizutani, K. 2017. Routing or Computing? The Paradigm Shift Towards Intelligent Computer Network Packet Transmission Based on Deep Learning. *IEEE Transactions on Computers*, 66(11): 1946–1960.

Nemirovski, A. 2003. On tractable approximations of randomly perturbed convex constraints. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, volume 3, 2419–2422. IEEE.

Sedghi, M.; Ahmadian, A.; and Aliakbar-Golkar, M. 2016. Assessment of optimization algorithms capability in distribution network planning: Review, comparison and modification techniques. *Renewable and Sustainable Energy Reviews*, 66: 415–434.

Shen, Y.; Dai, W.; and Win, M. Z. 2014. Power Optimization for Network Localization. *IEEE/ACM Transactions on Networking*, 22(4): 1337–1350.

Szabo, C.; Radenovic, V.; Judd, G.; Craggs, D.; Lee, K. L.; Chen, X.; and Chan, K. 2020. Optimizing Communication Strategies in Contested and Dynamic Environments. In *2020 25th International Conference on Engineering of Complex Computer Systems (ICECCS)*, 197–205.

Tu, K.; Chen, Z.; and Yue, M.-C. 2024. A Max-Min-Max Algorithm for Large-Scale Robust Optimization. arXiv:2404.05377.

Wang, M.; Cui, Y.; Wang, X.; Xiao, S.; and Jiang, J. 2018. Machine Learning for Networking: Workflow, Advances and Opportunities. *IEEE Network*, 32(2): 92–99.

Winstein, K.; and Balakrishnan, H. 2013. Tcp ex machina: Computer-generated congestion control. *ACM SIGCOMM Computer Communication Review*, 43(4): 123–134.

Yang, K.; Wu, Y.; Huang, J.; Wang, X.; and Verdú, S. 2008. Distributed robust optimization for communication networks. In *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, 1157–1165. IEEE.

Zhang, J.; Chen, X.; Xiang, Y.; Zhou, W.; and Wu, J. 2014. Robust network traffic classification. *IEEE/ACM transactions on networking*, 23(4): 1257–1270.