

Fang Fang
Fu Shu (Eds.)



457

LNICST

Game Theory for Networks

11th International EAI Conference, GameNets 2022
Virtual Event, July 7–8, 2022
Proceedings



Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering

457

Editorial Board Members

Ozgur Akan

Middle East Technical University, Ankara, Turkey

Paolo Bellavista

University of Bologna, Bologna, Italy

Jiannong Cao

Hong Kong Polytechnic University, Hong Kong, China

Geoffrey Coulson

Lancaster University, Lancaster, UK

Falko Dressler

University of Erlangen, Erlangen, Germany

Domenico Ferrari

Università Cattolica Piacenza, Piacenza, Italy

Mario Gerla

UCLA, Los Angeles, USA

Hisashi Kobayashi


Princeton University, Princeton, USA

Sergio Palazzo

University of Catania, Catania, Italy

Sartaj Sahni

University of Florida, Gainesville, USA

Xuemin Shen 

University of Waterloo, Waterloo, Canada

Mircea Stan

University of Virginia, Charlottesville, USA

Xiaohua Jia

City University of Hong Kong, Kowloon, Hong Kong

Albert Y. Zomaya

University of Sydney, Sydney, Australia

More information about this series at <https://link.springer.com/bookseries/8197>

Fang Fang · Fu Shu (Eds.)

Game Theory for Networks

11th International EAI Conference, GameNets 2022
Virtual Event, July 7–8, 2022
Proceedings

 Springer

Editors

Fang Fang
Western University
London, ON, Canada

Fu Shu
Chongqing University
Chongqing, China

ISSN 1867-8211 ISSN 1867-822X (electronic)
Lecture Notes of the Institute for Computer Sciences, Social Informatics
and Telecommunications Engineering
ISBN 978-3-031-23140-7 ISBN 978-3-031-23141-4 (eBook)
<https://doi.org/10.1007/978-3-031-23141-4>

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2022

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

We are delighted to introduce the proceedings of the 11th edition of the European Alliance for Innovation (EAI) International Conference on Game Theory for Networks (GameNets 2022). This conference brought together researchers, developers, and practitioners around the world who are leveraging and developing game theory for networks. The theme of GameNets 2022 was “game theory for next-generation networks”.

The technical program of GameNets 2022 consisted of 25 full papers that were selected from XX submissions in a XX blind review process, with each submission receiving at least XX reviews. The papers were distributed into the following conference tracks: Track 1 - Wireless Networks, Track 2 – Internet of Things, Track 3 – Game Theory. Aside from the high-quality technical paper presentations, the technical program also featured three keynote speeches given by Zhiguo Ding from the University of Manchester, UK, Haijun Zhang from the University of Science and Technology Beijing, China, and Yuan Wu from the University of Macau, Macao, China.

Coordination with the steering committee was essential for the success of the conference. We sincerely appreciate their constant support and guidance. It was also a great pleasure to work with such an excellent organizing committee team for their hard work in organizing and supporting the conference. In particular, we are grateful to the Technical Program Committee, who have completed the peer-review process for the technical papers and helped to put together a high-quality technical program. We are also grateful to Conference Manager Ivana Bujdakova for her support and to all the authors who submitted their papers to the GameNets 2022 conference.

We strongly believe that the GameNets conference provides a good forum for all researchers, developers, and practitioners to discuss all science and technology aspects that are relevant to game theory. We also expect that the future GameNets conferences will be as successful and stimulating as this year’s, as indicated by the contributions presented in this volume.

[AQ1](#)

Fang Fang
Shu Fu
Quoc-Viet Pham
Huan Zhou

Organization

Steering Committee

Imrich Chlamtac	University of Trento, Italy
Victor C. M. Leung	The University of British Columbia, Canada

Organizing Committee

General Chair

Fang Fang	Western University, Canada
-----------	----------------------------

General Co-chair

Fu Shu	Chongqing University, China
--------	-----------------------------

Technical Program Committee Chairs

Quoc-Viet Pham	Pusan National University, South Korea
Huan Zhou	Three Gorges University, China

Sponsorship and Exhibit Chair

Yueping Cai	Chongqing University, China
-------------	-----------------------------

Local Chair

Chen Chen	Chongqing University, China
-----------	-----------------------------

Workshops Chairs

Mingzhe Chen	Princeton University, USA
Zhaohui Yang	University College London, UK

Publicity and Social Media Chair

Lei Yuan	Lanzhou University, China
----------	---------------------------

Publications Chair

Donghong Cai	Jinan University, China
--------------	-------------------------

Web Chair

Peng Yu

Beijing University of Posts and
Telecommunications, China

Technical Program Committee

Aresh Dadlani

Nazarbayev University, Kazakhstan

Chao Song

University of Electronic Science and Technology
of China, China

Chi Lin

Dalian University of Technology, China

Dawei Li

Montclair State University, USA

Duong Tung Nguyen

Arizona State University, USA

Honglong Chen

China University of Petroleum, Qingdao, China

Md. Jalil Piran

Sejong University, South Korea

Liang Zhao

China Three Gorges University, China

Nhu Ngoc Dao

Sejong University, South Korea

Ning Wang

Rowan University, USA

Rui Chen

China Three Gorges University, China

Rukhsana Ruby

Shenzhen University, China

Wenjun Jiang

Hunan University, China

Xin Li

Nanjing University of Aeronautics and
Astronautics, Nanjing, China

Xuxun Liu

South China University of Technology,
Guangzhou, China

Zhaohui Yang

University College London, UK

Zhehui Xiong

Singapore University of Technology and Design,
Singapore

Contents

Wireless Networks

Block-Chain Abnormal Transaction Detection Method Based on Dynamic Graph Representation	3
<i>Chenbin Qiao, Yuanzheng Tong, Ao Xiong, Jing Huang, and Wei Wang</i>	
Multi-service Communication Isolation of Underground Pipe Gallery Based on WiFi6	16
<i>Shaojun Liu, Zheng Lou, Jingwei Li, Linlin Zhang, and Sujie Shao</i>	
MU-PDR: A Method of Fingerprint Passive Positioning for WiFi6 Based on MU-RTS/CTS	30
<i>Ningchi Zhang, Cheng Zhong, Qinghai Ou, Pengcheng Lu, La Bao, Yan Liu, and Sujie Shao</i>	
The Optimal Layer of User-Specific Reconfigurable Intelligent Surfaces Structure for Uplink Communication System	44
<i>Zihao Wen, Zhicheng Dong, Weixi Zhou, and Jing Zhao</i>	
A Sophisticated Anti-jamming Strategy for a Joint Radar and Communication System	55
<i>Andrey Garnaev and Wade Trappe</i>	

Internet of Things

Power Data Credible Decision-Making Mechanism Based on Federated Learning and Blockchain	73
<i>Xin Li, Fangjian Shang, Yanli Yao, and Tianren Zheng</i>	
A Service Protection Mechanism of Deterministic Networking Based on Segment Routing	83
<i>Tianchi Li and Yueping Cai</i>	
Cloud-Edge Collaboration Based Power IoT Scene Perception Mechanism	100
<i>Sujie Shao, Congzhang Shao, Cheng Zhong, Shaoyong Guo, and Pengcheng Lu</i>	
Task Allocation Mechanism of Power Internet of Things Based on Edge Routing Optimization	118
<i>Cheng Zhong, Yi Li, Sujie Shao, Shaoyong Guo, and Qianjun Wang</i>	

Voluntary Data Preservation Mechanism in Base Station-Less Sensor Networks 132
Yutian Chen, Jennifer Ly, and Bin Tang

Design and Implementation of Targeted Poverty Alleviation System Based on Blockchain Network 150
Jun Zhao, Wei Liang, Jingwen Li, Yan Liu, and Yan Liang

An Intent-Based Routing Scheme in Satellite IoT 157
Guanghui Zhang, Kang Liu, Xue Chen, Xu Liu, Zhen Pan, and Tianyang Jiang

Game Theory

Interference Management in Terrestrial-Satellite Networks Using Stackelberg Game 175
Yaomin Zhang, Haijun Zhang, and Keping Long

Optimal Resource Allocation for Computation Offloading in Maritime Communication Networks: An Energy-Efficient Design via Matching Game 187
Minghui Dai, Zhishen Luo, Tianshun Wang, Yuan Wu, Liping Qian, and Bin Lin

Game Theoretic Analysis of Resource Allocation in Multi-tiered Networks 200
Li Zhang, Mohit Hota, and Sanjiv Kapoor

Energy-Efficient Multi-cell NOMA Design via Coalition Formation Game 215
Yanqing Xu, Donghong Cai, Shuai Wang, Zhicheng Dong, and Weixi Zhou

The Vaccination Game in Susceptible-Infected-Susceptible (SIS) Networks with Multipopulations 230
Eitan Altman

A Stochastic Bandwidth Scanning Game 237
Andrey Garnaev and Wade Trappe

Port Capacity Leasing Games at Internet Exchange Points 251
Md Ibrahim Ibne Alam, Elliot Anshelevich, and Koushik Kar

A Phase Transition in Large Network Games 263
Abhishek Shende, Deepanshu Vasal, and Sriram Vishwanath

Dynamic Pricing for Tenants in an Automated Slicing Marketplace 278
Alessandro Lieto, Ilaria Malanchini, Silvio Mandelli, and Antonio Capone

Service Function Chain Placement in Cloud Data Center Networks: A Cooperative Multi-agent Reinforcement Learning Approach	291
<i>Lynn Gao, Yutian Chen, and Bin Tang</i>	
The Art of Concession in General Lotto Games	310
<i>Rahul Chandan, Keith Paarporn, Dan Kovenock, Mahnoosh Alizadeh, and Jason R. Marden</i>	
Providing Slowdown Information to Improve Selfish Routing	328
<i>Philip N. Brown</i>	
Budgeted Adversarial Network Resource Utilization Games	339
<i>Yi Zhang and Sanjiv Kapoor</i>	
Author Index	367

Author Queries

Chapter

Query Refs.	Details Required	Author's response
AQ1	Please provide the date at the end of the preface. And amend if necessary.	

Wireless Networks



Block-Chain Abnormal Transaction Detection Method Based on Dynamic Graph Representation

Chenbin Qiao¹(✉), Yuanzheng Tong¹, Ao Xiong¹, Jing Huang², and Wei Wang³

¹ Beijing University of Posts and Telecommunications, Beijing 100876, China
qiaochenbin@bupt.edu.cn

² Nanjing University of Aeronautics and Astronautics, Nanjing 200216, China

³ Beijing University of Technology, Beijing 100124, China

Abstract. The advent of cryptocurrency introduced by Bitcoin ignited an explosion of technological and entrepreneurial interest in payment processing. The user scale of Bitcoin is dynamic, and the participating identities are anonymous, which will lead to more hidden, sophisticated and intelligent money laundering crimes. Therefore, in order to realize intelligent anti-money laundering, it is necessary to accurately detect abnormal transactions. Recently, graph representation learning has shown strong advantages in the field of machine learning, and the current blockchain anomaly detection models based on graph representation learning are mainly designed for static graphs, however, real-world graphs evolve over time. Based on this, this paper proposes a block-chain abnormal transaction detection model DynAEGCN based on dynamic graph representation learning. This model uses the autoencoder as the framework. Firstly, the encoder uses the graph convolutional neural networks to gather neighborhood information to obtain low-dimensional feature vectors. Then, considering the dynamics of graphs, the GRU network is used to evolve the graph model itself over time. Finally, the decoder reconstructs the adjacency matrix and compares it with the real graph to construct the loss. Extensive experiments on the Bitcoin transaction dataset for edge classification tasks against financial crimes show that DynAEGCN model has better performance compared with related approaches.

Keywords: Dynamic graphs · Graphs convolution network · Graph representation learning

1 Introduction

The rapid development of science and technology has prompted the financial industry to move from the real economy to Internet finance, especially with the emergence of blockchain technology, virtual digital currencies represented by Bitcoin have poured into the financial market. With virtual currency, low-cost, peer-to-peer cross-border transactions can be realized. The use of blockchain to implement transactions has lowered the

entry threshold of the financial industry to a certain extent. Because the user participation, transaction and consensus mechanism of blockchain are open, the user scale is dynamic, and the participant identity is anonymous, which will lead to the more hidden process of money laundering crime, and the money laundering methods tend to be more complex and intelligent, thus bringing new problems and challenges to the traditional anti-money laundering supervision system.

Although it is difficult to achieve the supervision of blockchain transactions in traditional ways, the complete transaction data is open and transparent on the blockchain, mining the transaction data on the chain, establishing a multi-dimensional data model, and using technologies such as big data and artificial intelligence to achieve data-driven intelligent supervision schemes have become a new solution direction, and the primary goal of this intelligent supervision scheme is to accurately identify abnormal transactions on the blockchain: identify suspicious users (such as members of money laundering organizations) or suspicious transactions (such as credit card fraud transactions).

At present, representation learning based on graph structure data has become an important machine learning task, which is universally applicable in various structures such as social networks, cooperative networks, protein networks, and so on. Transactions on the blockchain can also be mapped as graph structure, the current blockchain anomaly detection model based on graph representation learning is mostly designed for static graphs, but transactions on the blockchain will change dynamically over time, as users continue to trade, the graph structure will also continue to change, so we not only need to pay attention to the information of the current moment in the graph, but also need to analyze the information of the historical moment on the graph.

Therefore, this paper combines the advantages of graph convolutional networks (GCN) in graph structure extraction, and the advantages of gated Recurrent Unit (GRU) in learning time series, and designs an anomaly detection model DyAEGCN learning graph structure and temporal characteristic information, which is used to mine the information in blockchain transaction data and realize a more essential portrayal of transaction data. Designed to improve abnormal transaction detection performance.

The DyAEGCN model takes the autoencoder as the framework. First of all, the encoder uses GCN to learn the structural characteristics of the network, aggregate the neighbor information of the nodes, and at the same time use the GRU adaptive update parameters to learn the time dynamics in the network, and finally build the loss in the comparison of the decoder reconstruction network with the real network. In this paper, the experiment is used to evaluate the node classification task, and compared with other algorithms, the experiment shows that the DyAEGCN model is better than the comparison algorithm in the edge classification task.

2 Related Work

The rapid development of blockchain technology has prompted the transformation of the Internet of Information to the Internet of Value, which has a wide range of application scenarios. However, due to the lack of regulatory mechanisms, many risks or violations of laws and regulations have been derived, such as money laundering, tax evasion and illegal ICO financing. Countries have incorporated blockchain technology

into their regulatory systems, and abnormal transaction detection has played a positive role in promoting the healthy development of the blockchain industry. The conventional solution for abnormal transaction detection is to design an alert system based on fixed threshold rules to detect and flag suspicious transactions, and then make human decisions or judgments on suspicious behaviors. However, the challenges faced by such regulatory schemes are reported: 1) How to construct effective rules from massive and heterogeneous transaction data, and keep the rules advanced and relevant; 2) How to set alarm thresholds for the calibration of suspicious transaction behaviors.

The emergence of Internet finance such as virtual “digital currencies” has created enormous challenges for rules-based regulatory solutions. It has become a trend to break the traditional supervision thinking and build an intelligent supervision scheme based on data and using technologies such as artificial intelligence and big data analysis. The research work on abnormal transaction detection in intelligent supervision mainly focuses on supervised learning and unsupervised learning. Learn. Supervised learning predicts the classification of unknown data samples (test set) by using labeled data (training set) to learn to discriminate binary (such as legal versus illegal transactions) or multi-class machine learning detection models. Unsupervised learning explores the structure and characteristics of unlabeled data, finds the optimal division of clusters or classes, and regards outliers far from other sample points as outliers, that is, abnormal data. For example, Jullum et al. [1] used information such as sender/receiver background, transaction early behavior, and transaction history to train an XGBoost supervised predictive model to identify potential money laundering behaviors in financial transactions and applied to banks. Paula et al. [2] extracted 18 important features from related categories such as registration information, financial transactions, and electronic invoices, and combined auto-encoder (AE, auto-encoder) algorithm to train an unsupervised deep learning model to detect and anti-money laundering-related export fraud;

At present, representation learning based on graph-structured data has become an important machine learning task. The basic idea of graph representation learning is to learn the low-dimensional vector representation of nodes, which requires the vector to retain the structural information and attribute information of nodes in the graph as much as possible etc., are generally applicable in various structures such as social networks, cooperative networks, protein networks, etc. Transactions on the blockchain can also be mapped to a financial network with users as nodes and transactions between users as edges. For example, Weber et al. [3] map Bitcoin transactions into a huge and complex graph structure, and extract the number of transactions and transaction amounts and other related features, and then use the graph convolution network (GCN, graph convolution network) algorithm to distinguish illegal and legal transactions.

At present, most static graph representation learning can effectively learn the vector representation of nodes, but a large amount of real data in life shows complex time characteristics, and transactions based on blockchain are also dynamic. The graph structure and its attributes will evolve dynamically over time. Nodes and edges in the graph will be inserted and deleted over time, and node attributes and edge attributes will also change over time. Therefore, we not only need to pay attention to the information of

the current moment in the graph, but also need to analyze the information of the historical moment on the graph. In this context, dynamic network modeling is important to accurately predict node attributes and future links.

As graph convolutional neural networks show great advantages in obtaining graph structure information, a new dynamic graph embedding learning method is to combine GCN and RNN, where GCN is used to extract information on graph structure, and RNN is used to extract information on graph structure. Used to model dynamic changes in the time dimension. Seo et al. [4] proposed two GCRN architectures, the common feature of which is to use GCN to learn the vector representation of nodes, and then input the vector sequence learned for a period of time into the LSTM model to model the dynamics on the sequence. The only difference between the two architectures is that one of the models modifies the Euclidean 2D convolution operation in the traditional LSTM to a graph convolution operation. Similarly, Manessi et al. [5] proposed WD-GCN/CD-GCN combining variants of LSTM and extended graph convolution operations to model graph structure and its long- and short-term dependencies. The difference is that the input of WD-GCN is a sequence of graphs, while the input of CD-GCN is an ordered sequence of corresponding node features. The Evolve-GCN model changes the idea of learning the dynamics of node representation in time series in the previous method, and instead learns the dynamics of GCN parameters in time series.

In view of the current situation and development trend of intelligent blockchain transaction supervision scheme, this paper takes full advantage of the advantages of graph convolutional neural network and recurrent neural network, and designs a dynamic graph representation learning model for mining blockchain transactions. The information in the data is designed to improve abnormal transaction detection performance.

3 The Structure of DynAEGCN

3.1 Problem Definition

We define dynamic graphs and representation learning for dynamic graphs as follows: a dynamic graph is represented as a sequence of multiple static graphs:

$$G = \{G_1, G_2, \dots, G_T\}, \quad (1)$$

Where $G_t = (V_t, E_t)$ denotes the snapshot at time t , $t \in \{1, 2, \dots, T\}$, The adjacency matrix of G_T is $A_t \in R^{N \times N}$. A node representation on a dynamic graph is learned as a sequence:

$$F = \{f_1, f_2, \dots, f_T\}, \forall t \in \{1, 2, \dots, T\}, \quad (2)$$

Each of the mappings maps the nodes to a low-dimensional vector $(y_t)_v = f_t(v)$, so that the mapped vector can retain the original information of the node. That is to say, the more similar two points are in the original image, the closer their mapped vectors are.

We consider a multi-layer Graph Convolutional Network (GCN) with the following layer-wise propagation rule:

$$H^{l+1} = \sigma \left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right), \quad (3)$$

Here, $\hat{A} = A + I$, I is the adjacency matrix of the undirected graph G with added self-connections. $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$, \hat{D} is the degree matrix of \hat{A} ; the operation $\hat{D}_t^{-\frac{1}{2}} \hat{A}_t \hat{D}_t^{-\frac{1}{2}}$ is a heterogeneous normalization of the adjacency matrix as an approximate graph Convolution filter; W_l^l is the weight matrix of the l th layer; $\sigma(\cdot)$ denotes an activation function. The input X_t^0 of the first layer of the network is the feature matrix of the node, and each row of the matrix is the K -dimensional feature vector of each node.

3.2 The Architecture of DynAEGCN

This paper aims to solve the problems of complexity and dynamics in the dynamic transaction network of the blockchain. The model proposed in this paper adopts the classic unsupervised self-encoding framework learning, which uses the encoder to encode the input graph A to generate the feature X , and the feature X is generated using the decoder, by minimizing the distance between A and A' , allowing the decoder to learn the ability to predict the graph while the encoder maps the input graph to a vector space (Fig. 1).

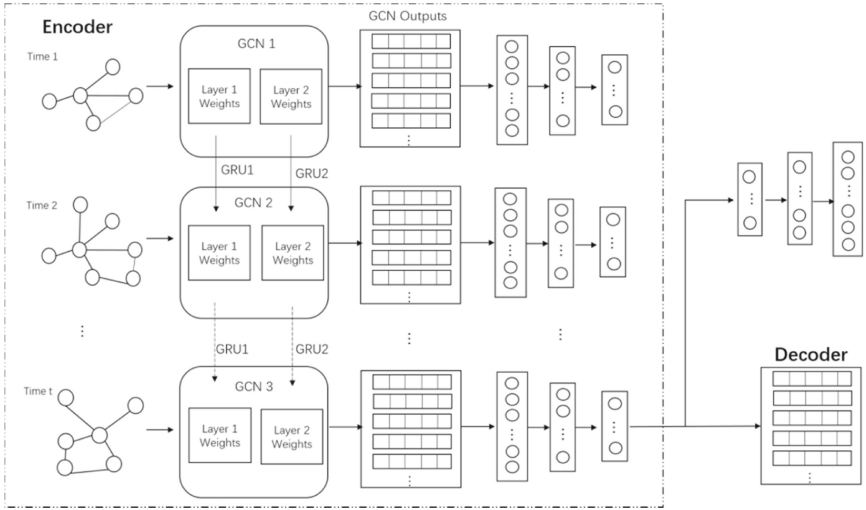


Fig. 1. The Architecture of DynAEGCN

- 1) In the encoder part, the graph convolutional network (GCN) model is adopted in the time dimension, and the structure and temporal feature information of the graph are learned by using the RNN to evolve the GCN parameters, in which the GCN adopts a two-layer neural network. This approach effectively performs model tuning, which focuses on the model itself rather than node embeddings. Therefore, there is no limit to node changes. Furthermore, for future graphs with new nodes without historical information, the evolved GCN is still plausible for them.
- 2) The decoder can reconstruct the original graph and compare it with the original graph to construct the loss, so that the link relationship between nodes can be learned more accurately.

3.3 Encoder

In the encoder part, we take advantage of the graph structure extraction advantage of GCN to learn the structural information under each time slice. GCN aggregates neighbor information through the defined spectral graph convolution, thereby extending the idea of convolution to the graph. Formally, for $G_t = (V_t, E_t)$ at time t, the input of the lth layer of GCN is the vector X_t^l and the adjacency matrix A_t output by the $l - 1$ th layer, and the output is the updated node vector X_t^{l+1} . The operation at the l level is expressed as:

$$X_t^{l+1} = F(X_t^l, A_t, W_t^l) = \sigma \left(\hat{D}_t^{-\frac{1}{2}} \hat{A}_t \hat{D}_t^{-\frac{1}{2}} X_t^l W_t^l \right), \quad (4)$$

Among them, the superscript l represents the l -th convolutional layer, and the subscript t represents the t -th time step; $\hat{A}_t = A_t + I$, I is the identity matrix; $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$, \hat{D} is the degree matrix of \hat{A} ; the operation $\hat{D}_t^{-\frac{1}{2}} \hat{A}_t \hat{D}_t^{-\frac{1}{2}}$ is a heterogeneous normalization of the adjacency matrix as an approximate graph Convolution filter; W_t^l is the weight matrix of the l th layer at time t ; σ is the nonlinear activation function. The input X_t^0 of the first layer of the network is the feature matrix of the node at time t , and each row of the matrix is the K -dimensional feature vector of each node. After the graph convolution operation of L layers, the neighbor information of the node is aggregated in the output vector of each time slice.

Considering the dynamic nature of the graph, the dynamic convolution layer adds an update mechanism to the static GCN architecture. Because when the graph structure changes, the weight parameters of the convolution operation should also be updated in order to adapt to the new graph structure. Recurrent Neural Network (RNN) is a type of recurrent neural network that takes sequence data as input, performs recursion in the evolution direction of the sequence, and connects all nodes (recurrent units) in a chain. Recurrent neural network has memory, parameter sharing and Turing completeness, so it has certain advantages in learning nonlinear characteristics of sequences.

In this paper, the RNN component is used to update the weight parameters of the GCN model. For each $t \in \{1, 2, \dots, T\}$ and $l \in \{1, 2, \dots, L\}$, the RNN components use the parameters of The initial value is used as input, and the updated W_t^l is output. Gated Recurrent Unit (GRU) is a variant of Recurrent Neural Network (RNN). Since RNN has the problem of gradient dispersion and gradient explosion, it is often far from the expected effect, so the GRU network is proposed. RNN and GRU networks are also modeled using the previous hidden state and the current input, the difference is that the latter uses reset gates and update gates in the internal structure of the device. GRU can introduce richer graph structure information for the update of weight parameters, so our architecture adopts the implementation of GRU. The weight update method of the l th layer at time t is as follows:

$$W_t^l = G(X_t^l, W_{t-1}^l) = (1 - Z_t^l) \circ W_{t-1}^l + Z_t^l \circ \hat{W}_t^l, \quad (5)$$

$$Z_t^l = \sigma(U_Z^l X_t^l + V_Z^l W_{t-1}^l + B_Z^l), \quad (6)$$

$$R_t^l = \sigma(U_R^l X_t^l + V_R^l W_{t-1}^l + B_R^l), \quad (7)$$

$$\hat{W}_t^l = \tanh(U_W^l X_t^l + V_W^l (R_t^l \circ W_{t-1}^l) + B_W^l), \quad (8)$$

Among them, Z_t^l , R_t^l , \hat{W}_t^l are the update gate output, reset gate output and pre-output, respectively.

The update of the weight matrix can be seen as applying standard GRU operations to each column of the matrix. The standard GRU operation is for between vectors, and the process of updating the GCN weight matrix is for the operation between matrices. The weight matrix \hat{W}_t^l at time t is used as the hidden state of the GRU; the node representation matrix X_t^l at the l th layer at time t is used as the input of the GRU unit to introduce the information of the current time; the GRU unit outputs the updated W_{t+1}^l , as the weight matrix at the next moment. The calculation of W_{t+1}^l includes the information of the historical moment and the current moment. Since the weight matrix \hat{W}_t^l and the node representation matrix X_t^l have different column dimensions, the sampling of X_t^l is newly added to the network layer operation of this layer to achieve the same number of columns as \hat{W}_t^l Scheduling Method.

The GCN module aggregates the neighbor information of nodes, while the GRU updates the weight parameters with the time dimension. That is, the encoder can be expressed as:

$$X_t^{l+1} = F(X_t^l, A_t, W_t^l) = F(X_t^l, A_t, G(X_t^l, W_{t-1}^l)), \quad (9)$$

3.4 Decoder

The decoder reconstructs the adjacency matrix from the information of the first t time steps learned by the encoder, which is the topological map of the predicted time $t + 1$. The decoder uses the dot product to reconstruct the original image, and the decoding process is expressed as:

$$\hat{A} = \sigma(ZZ^T), \quad (10)$$

where \hat{A} denotes the reconstructed adjacency matrix.

The adjacency matrix directly determines the topology of the graph, so the goal of this model is to make the reconstructed adjacency matrix as similar to the original adjacency matrix as possible, compare the two to construct a loss, and backpropagate the updated parameters to learn the representation of the hidden layer nodes.

3.5 Loss Function

To test the representation ability of the model, we train the model on a specific edge classification task. The task of edge classification has strong practical significance in many real-world scenarios. For example, to identify crimes in financial networks, it is necessary to conduct edge classification research on the connection between two accounts. The edge classification task under dynamic graph aims to predict the edge label category of an edge (u, v) at time t . To classify an edge, we need the node vector

representation of the two endpoints of the edge. Given a vector representation of two nodes u and v connected by an edge at time t . Given that the vector representations of two nodes u and v connected by an edge at time t are X_t^u and X_t^v respectively, the parameter matrix P is used to predict the label of the edge (u, v) :

$$y_t^{uv} = \text{softmax}(P[X_t^u, X_t^v]), \quad (11)$$

The cross-entropy loss function of the model is:

$$L = - \sum_{t=1}^T \sum_{(u,v)} \alpha_{uv} \sum_{i=1}^N (Z_t^{uv}) \log(y_t^{uv})_i, \quad (12)$$

Among them, Z_t^{uv} represents the true label category of the edge; the weight parameter α_{uv} is a hyperparameter for balancing the weight of the category distribution. The experimental datasets all have serious category imbalance problems, and the proportion of classification categories is balanced by adjusting α_{uv} .

4 Simulation Analysis

4.1 Datasets

Model validation is performed on two blockchain finance domain datasets from a trust scoring network between users of two different bitcoin trading websites. The Bitcoin OTC dataset is a network of trust scores among users extracted from Bitcoin trading websites. Users rate other users from -10 (completely distrusted) to $+10$ (completely trusted), and each rating has a corresponding timestamp representing the scoring time. The time span of the dataset is about 5 years, we set a time interval of 13.8 days, and the dataset produces a total of 138 time steps. The 138 time steps are split into training, validation and test sets. The category distribution of the Bitcoin OTC dataset is extremely uneven, 89% of the data are positive examples, and negative examples only account for a very small part.

The Bitcoin Alpha dataset is also a network of trust among Bitcoin users, but the user and rating data are drawn from another Bitcoin platform, BTC-Alpha. The scoring data is from November 8, 2010 to January 22, 2016, with a time interval of 13.6 days, and the dataset is divided into 140 time steps. Scores still range from -10 (complete distrust) to $+10$ (complete trust), with Bitcoin Alpha having a higher positive ratio (93%) than Bitcoin OTC.

4.2 Contrast Models

We compare the DynAEGCN model with some existing static and dynamic methods.

GCN is a static graph convolution model and a classical method for graph representation learning. The model uses spectral convolution to aggregate the neighbor information of nodes to learn the embedding vector of nodes. Because each time step in the dynamic graph will produce a snapshot of the graph, we use the same GCN model for each time

step, that is, regardless of the dynamics of the graph, the GCN model trains the graph on each time step.

GCN-GRU combines GCN with sequence modeling. First, the representation vectors of nodes under each time snapshot are learned through GCN architecture, and then these vectors are input into GRU unit to learn the dynamics of node representation. The dynamic representation of this method is based on the node representation vector, which belongs to the node oriented method.

EvolveGCN is a model oriented approach. This method also combines GCN with RNN, but different from GCN-GRU, RNN is used in evolving GCN to model the update of GCN parameters. The whole model training is carried out from bottom to top along the convolution layer and from front to back in time dimension. The dynamics is modeled into the implicit vector of RNN. For each time step, the updated GCN weight parameters are learned, and the graph convolution operation is carried out to obtain the updated node representation.

4.3 Experimental Results

The table shows the comparison results of the classification performance of the DynAEGCN model and the comparison model. The unbalanced categories of the two datasets make the model classification ability face great challenges, but the DynAEGCN model in this paper achieves the best classification ability and has obvious advantages over other models. For the overall classification ability, we compare the accuracy rate and the weighted F1 value; since the subclass has stronger practical significance for anti-financial fraud, we also compare the F1 value of the subclass and the corresponding precision and recall rate (Table 1).

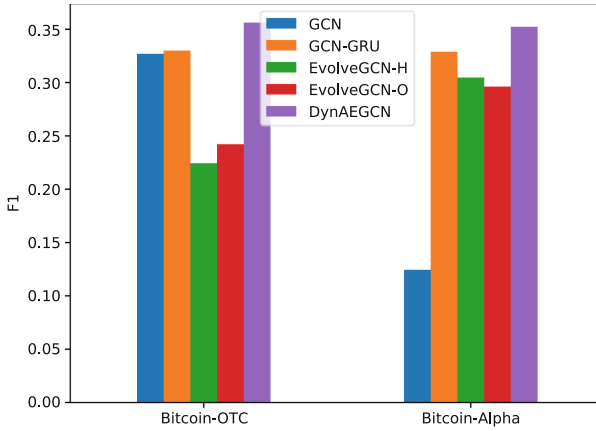
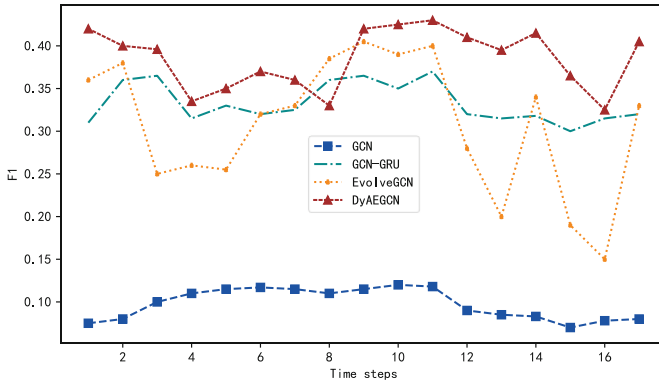
Table 1. Experimental results for edge classification tasks on the Bitcoin OTC dataset

Method	Accuracy	Precision	Recall	F1-score
GCN	56.55%	21.05%	73.18%	32.69%
GCN-GRU	60.71%	22.41%	62.47%	32.99%
EvolveGCN-H	67.40%	17.84%	30.19%	22.43%
EvolveGCN-O	68.31%	18.95%	33.52%	24.21%
DynAEGCN	74.42%	32.01%	40.15%	35.62%

It can be found from Table 2 that DynAEGCN has the highest classification accuracy and weighted F1 value, indicating that DynAEGCN has good overall classification performance. For the classification results of small classes, DynAEGCN also achieves the best F1 value and accuracy. Although the recall rate of DynAEGCN is slightly lower than that of GCN and GCN-GRU, DynAEGCN is still better than other methods because other methods have higher recall rate but lower accuracy rate. As the harmonic average of accuracy and recall, F1 is a more effective evaluation standard in classification performance (Figs. 2 and 3).

Table 2. Experimental results for edge classification tasks on the Bitcoin Alpha dataset

Method	Accuracy	Precision	Recall	F1-score
GCN	65.13%	11.50%	13.55%	12.44%
GCN-GRU	42.68%	22.62%	60.18%	32.88%
EvolveGCN-H	73.39%	25.34%	38.19%	30.47%
EvolveGCN-O	67.73%	23.03%	41.48%	29.62%
DynAEGCN	79.26%	33.54%	37.11%	35.23%

**Fig. 2.** Performance of edge classification**Fig. 3.** F1 score over time

Furthermore, we plot the F1 value and classification accuracy over time on the test set. It can be seen that the static GCN method is obviously different from other dynamic methods, and the advantages of DynAEGCN are more obvious in each time step. In

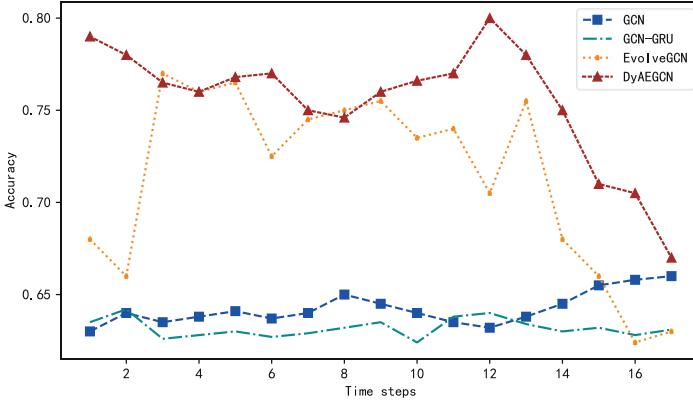


Fig. 4. Accuracy score over time

addition, the accuracy of GCN method is also lower. Because GCN is designed for static graphs and does not consider the dynamics of graphs, the performance disadvantages of GCN on dynamic graphs reflect the necessity and advantages of dynamic modeling. As can be seen from Fig. 4, the advantages of the DynAEGCN model can be maintained throughout the time axis. In particular, for time step 15, the classification ability of other methods is very poor, while the DynAEGCN model still retains the absolute advantage of F1 value. This is due to the dual modeling of spatiotemporal information by DynAEGCN model, which can have a relatively stable performance for abrupt changes in time series. In addition, compared with the two types of EvolveGCN, DynAEGCN has better classification performance, because EvolveGCN only focuses on the dynamics of model weight parameters and ignores the changes of graph structure. GCN-GRU has a relatively lower classification ability because although the historical information of each node is considered, DynAEGCN is still more advantageous for higher-level representation learning due to DynAEGCN's unique spatio-temporal convolution operation and model update mechanism.

5 Conclusion

In this paper, a dynamic graph representation learning model DynAEGCN is proposed to mine the implicit relationship between blockchain transaction features. The advantage of the recurrent unit in learning time series, the information of the time dimension and the space dimension can be aggregated, and more effective node representation can be learned. Further, the edge classification task is performed in a dynamic financial network with extremely imbalanced classification of two categories, and the results show that the DynAEGCN model outperforms all the contrasting models. The research on dynamic graph representation learning in this paper has strong practical significance, and can also provide a variety of possibilities for future research directions. In the follow-up work, the scalability of the model can be further improved, and the graph representation learning task of the model can be extended to a wider range of fields, such as node classification,

link prediction and clustering, etc., while increasing the learning and analysis of datasets in other fields.

Funding Statement. This work was supported by National Key R&D Program of China under Grant 2020YFB1005900.

References

1. Weber, M., Domeniconi, G., Chen, J., et al.: Anti-money laundering in bitcoin: experimenting with graph convolutional networks for financial forensics. arXiv Preprint arXiv:1908.02591 (2019)
2. Jullum, M., Løland, A., Huseby, R.B., et al.: Detecting money laundering transactions with machine learning. *J. Money Laund. Control* **23**(1), 173–186 (2020)
3. Paula, E.L., Ladeira, M., Carvalho, R.N., et al.: Deep learning anomaly detection as support fraud investigation in Brazilian exports and anti-money laundering. In: 2016 15th IEEE International Conference on Machine Learning and Applications, pp. 954–960. IEEE Press, Piscataway (2016)
4. Seo, Y., Defferrard, M., Vandergheynst, P., et al.: Structured sequence modeling with graph convolutional recurrent networks [EB/OL], 22 December 2016. Accessed 20 May 2020
5. Manessi, F., Rozza, A., Manzo, M.: Dynamic graph convolutional networks. *Pattern Recogn.* **97**, 107000 (2020)
6. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks [EB/OL], 22 February 2017. <https://arxiv.org/abs/1609.02907>. Accessed 20 May 2020
7. Chen, J., Ma, T., Xiao, C.: FastGCN: fast learning with graph convolutional networks via importance sampling [EB/OL], 30 January 2018. <https://arxiv.org/abs/1801.10247>. Accessed 20 May 2020
8. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: Guyon, I., Luxburg, U.V., Bengio, S., et al. (eds.) *Advances in Neural Information Processing Systems*, pp. 1024–1034. Current Associates, Long Beach (2017)
9. Yang, Y., Wu, Q.M.J., Wang, Y.: Autoencoder with invertible functions for dimension reduction and image reconstruction. *IEEE Trans. Syst.* **48**(7), 1065–1079 (2016)
10. Perozzi, B., Al-Rfou, R., Skiena, S.: DeepWalk: online learning of social representations. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 701–710. ACM, New York (2014)
11. Panzarasa, P., Opsahl, T., Carley, K.M.: Patterns and dynamics of users’ behavior and interaction: network analysis of an online community. *J. Am. Soc. Inform. Sci. Technol.* **60**(5), 911–932 (2009)
12. Hamilton, W.L., Ying, R., Leskovec, J.: Inductive representation learning on large graphs. In: *NIPS 2017: Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 1025–1035. ACM, New York (2017)
13. Li, J., Dani, H., Hu, X., et al.: Attributed network embedding for learning in a dynamic environment. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, New York, pp. 387–396 (2017)
14. Du, L., Wang, Y., Song, G., et al.: Dynamic network embedding: an extended approach for skip-gram based network embedding. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, Stockholm, pp. 2086–2092 (2018)
15. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 855–864. ACM, New York (2016)

16. Trivedi, R., Dai, H., Wang, Y., et al. Know-evolve: deep temporal reasoning for dynamic knowledge graphs. In: Proceedings of the 34th International Conference on Machine Learning, pp. 3462–3471. ACM, New York (2017)
17. Tang, J., Qu, M., Wang, M., et al.: Line: large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web, Florence, pp. 1067–1077 (2015)
18. Nguyen, G.H., Lee, J.B., Rossi, R.A., et al.: Continuous-time dynamic network embeddings. In: Companion Proceedings of the Web Conference, Lyon, pp. 969–976 (2018)
19. Wang, D., Cui, P., Zhu, W.: Structural deep network embedding. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, p. 1225 (2016)



Multi-service Communication Isolation of Underground Pipe Gallery Based on WiFi6

Shaojun Liu¹, Zheng Lou¹, Jingwei Li¹, Linlin Zhang^{2(✉)}, and Sujie Shao²

¹ Nanjing Power Supply Branch, State Grid Jiangsu Electric Power Co., Ltd., Nanjing, China

² State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China
{zll1618, buptssj}@bupt.edu.cn

Abstract. There are a variety of services in underground pipe gallery, and these services should be isolated during data transmission to ensure communication security. WiFi6(802.11ax) has the advantages of fast transmission rate, low energy consumption and low cost, and is suitable for carrying the underground pipe gallery communication service. In order to ensure the reliability of underground pipe gallery service transmission, it is necessary to study the communication isolation technology of WiFi6. In this paper, a communication isolation technology based on WiFi6 is proposed, and the communication isolation architecture of underground pipe gallery is designed to realize communication service isolation in frequency domain and spatial domain. In the frequency domain, a spectrum isolation scheme based on guardband is designed. In the spatial domain, a spatial stream isolation scheme based on service grouping is designed. Then, an isolation resource allocation method based on improved KM algorithm is proposed to realize the optimization of multi-service isolation. The simulation results show that the proposed communication isolation scheme can realize the communication isolation of underground pipe gallery with lower resource cost.

Keywords: Communication service isolation · WiFi6 · Underground pipe gallery · Guardband · KM algorithm

1 Introduction

The underground pipe gallery contains lines and pipes of different systems, such as water, electricity and gas, which share the space of the underground pipe gallery and have various communication demands [1]. In underground pipe gallery, data interaction between multiple types of services is generally achieved by wireless communication, which reduces wiring costs [2]. WiFi6(802.11ax), as the latest generation of IEEE 802.11 protocol that can be compared with 5G in terms of communication capability [3], becomes a technology selection suitable for multi-service communication in underground pipe gallery. In addition, multiple services in the underground pipe gallery should be isolated to ensure that communication services with different requirements [4], such as production, scheduling, management, and administration, do not affect each other.

Therefore, it is necessary to design a proper communication isolation scheme according to the service characteristics of underground pipe gallery and the technical characteristics of WiFi6.

Compared with traditional WiFi technology, WiFi6 has great progress in energy consumption, data transmission speed and multi-service carrying capacity [5, 6]. OFDMA (Orthogonal Frequency Division Multiple Access) and MU-MIMO (Multi-User Multiple-Input Multiple-Output) are two important technologies in WiFi6. They can physically divide communication resources and play a role in communication service isolation. OFDMA is a frequency division multiplexing technology, which divides the bandwidth into multiple non-overlapping subcarriers [7]. The subcarrier bandwidth in WiFi6 is defined as 78.125 kHz, and several OFDMA subcarriers can form a RU (Resource Unit). The orthogonality between subcarriers ensures that there will be no interference in signal transmission, thus ensuring the isolation of subcarriers and achieving physical isolation effect [8]. MU-MIMO technology allows multiple users to access the channel at the same time through physical multi-antenna [9]. Due to the spatial multi-path propagation characteristics, when the spacing between each antenna unit is large enough, the multipath fading between receiver and transmitter antennas tends to be independent, and the service transmission channel also tends to be independent [10], realizing isolation in the spatial transmission process. WiFi6 can support OFDMA and MU-MIMO at the same time. When the number of subcarriers contained in RU is greater than or equal to 106, it can support MU-MIMO [11].

In this paper, a multi-service communication isolation scheme of underground pipe gallery based on WiFi6 is designed by combining the characteristics of underground pipe gallery and WiFi6. Firstly, an underground pipe gallery communication multi-service isolation architecture is designed to achieve service isolation in frequency domain and spatial domain. After that, the physical isolation method of multi-service in underground pipe gallery is designed. The service isolation method based on guardband is proposed in frequency domain, and the service isolation method based on service grouping is proposed in spatial domain. In the resource allocation stage, the mapping relationship between services and resources is established by using the algorithm based on graph theory. Finally, the feasibility of the communication isolation scheme is verified by simulation.

2 Multi-service Communication Isolation Architecture

In view of the technical characteristics of WiFi6 multi-service transmission and the isolation requirements of different communication services in underground pipe gallery, this paper proposes a communication multi-service isolation architecture based on WiFi6, which realizes multi-service isolation from the perspectives of frequency domain and spatial domain. The specific architecture is shown in Fig. 1. The isolation architecture consists of three layers: service layer, frequency domain layer and spatial domain layer, which bear different functions respectively.

In the service layer, the communication requests sent by each terminal are summarized and the communication services are classified according to the differences in communication service resource requirements and isolation requirements. After that,

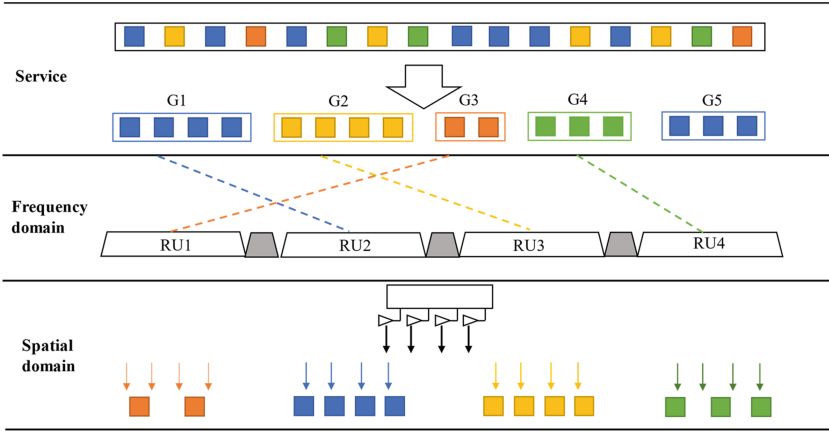


Fig. 1. Underground pipe gallery communication service isolation architecture.

communication services are divided into multiple service groups based on service types for subsequent communication isolation and resource allocation.

In the frequency domain layer, a spectrum isolation scheme based on the guardband is designed to realize the isolation of services in the frequency domain by limiting the distance of RU carrying services in the spectrum. According to the isolation requirements of service group and the guardband provided by RU, the mapping relationship between service group and RU is established based on the idea of graph theory. The RU specification of frequency domain division is 106-tone RU, which is the minimum RU that can support MU-MIMO, facilitating communication resource reuse and secondary isolation in spatial domain.

In the spatial domain layer, a spatial stream isolation scheme based on service grouping is designed. After the service group obtains RU, the spatial stream resources in the spatial domain are allocated among the services within the service group. Since different services are transmitted by different spatial streams in different spatial paths, the spatial stream isolation is realized.

In the isolation architecture, when a terminal sends a communication service request, the service type will be determined at the service layer first, and it will form a service group with several similar services. Secondly, in the frequency domain layer, the service groups compete for RU resources on the spectrum. At this time, all the service in the same RU is of the same kind of service, and the isolation of different kinds of service is realized in the frequency domain. Then, after the service group successfully obtains RU, the spatial stream under the RU is allocated among the services within the service group. The services obtain different spatial streams and transmit them in different paths, and realize the services isolation within a service group in the spatial domain. Finally, the resource allocation scheme is issued, and the communication service is designated to in certain spatial streams under a certain RU or is notified that the resource request fails. In this way, communication services can be isolated in both frequency domain and spatial domain, and different isolation requirements of communication services can be taken into account.

3 Multi-service Communication Isolation Method for Underground Pipe Gallery Based on WiFi6

3.1 Spectrum Isolation Scheme Based on Guardband

Isolation Mechanism Based on Guardband. The guardband is used to enhance the isolation in the frequency domain. It treats each RU as an independent channel and achieves isolation by limiting the distance between RU that transmitting services.

The guardband is the empty spectrum between channels. Frequency division multiplexing technology modulated different services to different spectrum positions. When the spectrum distance between services is wide enough, it can ensure that the signals do not interfere with each other [12]. OFDMA technology in WiFi6 divides the working bandwidth into several subcarriers, and the orthogonality of subcarriers provides natural isolation in frequency domain. There are Pilot subcarriers, DC subcarriers, Guard subcarriers and Null subcarriers on the channel [13]. The DC subcarriers, Guard subcarriers and Null subcarriers are not used for data transmission, but are used to resist interference from adjacent channels or adjacent subchannels. Figure 2 shows the division scheme of 106-tone RU at 20 MHz/40 MHz/80 MHz working bandwidth, while the null subcarriers and pilot subcarriers are not shown in the figure. When all RU on the spectrum carry data, some RU are too close to each other on the spectrum to meet the high isolation requirements of some services. Therefore, guardband can restrict the spectrum distance between RU carrying different services to ensure service isolation.

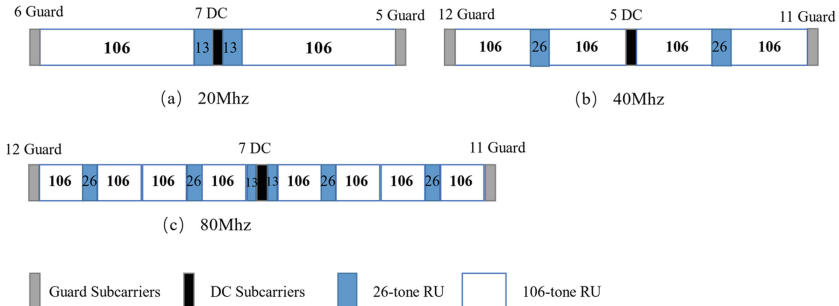


Fig. 2. 106-tone RU at 20 MHz/40 MHz/80 MHz.

According to the channel characteristics of WiFi6, when the guardband mechanism is adopted, part of the subcarriers originally used for data transmission will also be used for spectrum isolation due to the protection requirements of services, and part of communication resources will be sacrificed to enhance the isolation in the frequency domain. Multiple services of the same type form a service group. When service group applies for RU resource in the frequency domain, the guardband of the service group must be the same as that of the services in the service group. Figure 3 shows the allocation method of RU resources in the frequency domain at a certain time. Service can only be transmitted on the RU whose guardband is greater than the isolation requirement, and different types of services are in different RU.

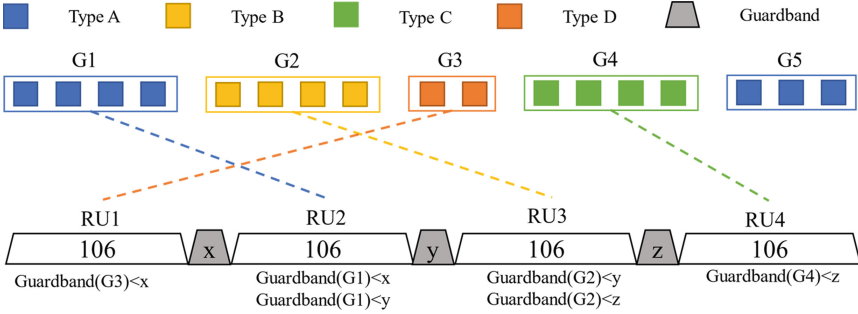


Fig. 3. Spectrum isolation mechanism based on guardband.

Guardband Calculation Method. WiFi6 guardband calculation mainly considers WiFi6 spectrum characteristics and service isolation requirements. WiFi6 and 5G both use OFDMA to divide spectrum resources, and some frequency bands in 5G are close to WiFi6 frequency band [14], so 5G guardband has strong reference for WiFi6 guardband.

The calculation formula of the guardband in 5G is Formula (1), Where $GB_{5G}(SCS, BW_{Band})$ is guardband. BW_{Band} is the working bandwidth, SCS is the sub-carrier space, NRB is the number of RB, and 12 consecutive subcarriers are defined as 1 RB.

$$GB_{5G}(SCS, BW_{Band}) = \frac{BW_{Band} - SCS * NRB * 12}{2} - \frac{SCS}{2} \quad (1)$$

Table 1. Guardband corresponding to different SCS and BW values in 5G.

SCS	BW = 5 MHz	BW = 10 MHz	BW = 20 MHz	BW = 30 MHz	BW = 40 MHz
15 kHz	242.5 kHz	312.5 kHz	452.5 kHz	592.5 kHz	552.5 kHz
30 kHz	505 kHz	665 kHz	805 kHz	945 kHz	905 kHz
60 kHz	N/A	1010 kHz	1330 kHz	1290 kHz	1610 kHz

Table 1 shows the guardband corresponding to some working bandwidths in 5G. Under the same working bandwidth, the value of the GB increases with the value of the SCS [15], while the value of the $\frac{GB(SCS)}{SCS}$ shows a downward trend on the whole. In WiFi6, the SCS is 78.125 kHz, and the $\frac{GB(60 \text{ KHz})}{60 \text{ KHz}} > \frac{GB(78.125 \text{ KHz})}{78.125 \text{ KHz}}$ can be judged to be true by the changing trend of 5G guardband.

$$GB_{RU}(RUband, \sigma_{Service}) = GB_{5G}\left(60, BW_{\lceil \frac{RUband}{5} \rceil * 5}\right) * \frac{78.125}{60} * \sigma_{Service} \quad (2)$$

The calculation method of RU guardband in WiFi6 is formula (2), the formula calculates a reference value according to 5G Guardband and introduces $\sigma_{Service}$. $\sigma_{Service}$ is the service impact factor, and its value is related to service isolation requirement.

When the total bandwidth of 106-tone RU is substituted, $GB_{RU} (78.125 \text{ KHz} * 106) = 1315 \text{ KHz} * \sigma_{Service}$. Because RU is only a sub-channel of WiFi6 working bandwidth, its environmental interference is far less than that of the complete channel with the same bandwidth. The bandwidth of a normal 20 MHz WiFi6 working channel edge Guard sub-carriers is about 500 kHz, and the reference value calculated by the above method exceeds the isolation requirements of normal services. To save precious spectrum resources, the value of $\sigma_{Service}$ is usually less than 1.

Table 2. Service classification and their isolation requirements in underground pipe gallery.

Service type	Isolation requirement	$\sigma_{Service}$	Transmission rate	Latency requirement	Typical service
Data collection	Low	0.1	≤ 10 Mbps	Low	Temperature sensing
Video and image	Medium	0.4	5–100 Mbps	Medium	Machine room monitoring
Voice call	High	0.7	≤ 2 Mbps	High	Real-time call
Emergency	Highest	1.0	2–100 Mbps	Highest	Fire alarm

Table 2 shows the service classification and their isolation requirements in underground pipe gallery. Put $\sigma_{Service}$ of four types of underground pipe gallery communication services into Formula (2), and calculate that the guardband of data collection services is 131.5 kHz, that of video image services is 526 kHz, and that of voice call services is 920.5 kHz. The guardband of emergency services is 1315 kHz.

3.2 Spatial Stream Isolation Scheme Based on Service Grouping

In the spatial domain, this paper uses the characteristics of WiFi6 multi-antenna to carry different services under the same service group with different spatial streams to realize the isolation of services that have the same isolation requirements.

When the service group successfully obtains RU, the spatial stream resources obtained by each RU through spatial division multiplexing are allocated within the service group. The division of the service group will affect the utilization rate of spatial resources. When there are too many services in a service group, some services in the service group cannot obtain spatial stream resources. When there are too few services in a service group, spatial resources are wasted. Therefore, service groups need to be divided based on the number of spatial resources of the WiFi6 device. Set the number of spatial streams supported by the WiFi6 device is N_{stream} , the service group division method is as follows:

1. **Statistical Service Information.** The number of services participating in this round of resource allocation is M , and the service is $V = \{V_1, V_2, \dots, V_M\}$, the transmission rate requirement is $R = \{R_1, R_2, \dots, R_M\}$, service group number X , X is

initialized to 1. The transmission rate of RU is B , and its calculation method is formula (3). SBL is the bit of each Symbol, $Code_{rate}$ is the bit rate, $N_{Subcarrier}$ is the number of subcarriers, N_{MIMO} is the number of spatial streams, Symbol is the signal transmission time in the time domain, a fixed value of 12.8 us, GI is the gap of inter-frame, β is the ratio of the actual transmission rate to the theoretical transmission rate in the current communication environment. When calculating the RU transmission rate of 106-tone RU, $N_{Subcarrier}$ is set to 106 and N_{MIMO} is set to 1.

$$B = (SBL \times Code_{rate} \times N_{Subcarrier} \times N_{MIMO}) \div (Symbol + GI) * \beta \quad (3)$$

2. **Create a Service Group G_X .** Set the value of $Num_X = 0$, indicates the number of allocated spatial streams. Select a service that is not added to the service group and select its service type K as the service type to be received by service group G_X .
3. **Add service to G_X .** All services that meet $V_i = k$ are sorted by the waiting time. The services with long waiting time are placed at the front of the sequence and then traversed in sequence. Calculates the number of spatial streams R_i/B required by service according to its transmission rate requirement. If $Num_X + R_i/B \leq N_{stream}$, the service is added successfully. When $Num_{mimo} + R_i/B > N_{stream}$, the join fails. G_X does not add new service when either of the following conditions is met: $Num_{mimo} = 8$ or the services that meet $V_i = k$ have been traversed. If all services have been added to the service group, the division of service groups is complete, and the value of X is the number of service groups. Otherwise, $X = X + 1$, to Step2.

When a service group obtains a RU in the frequency domain, RU obtains multiple spatial streams through MU-MIMO technology space division multiplexing, and these spatial streams are distributed among the services in the service group. According to the division method of service groups, the spatial stream resources under each RU can meet the transmission requirements of all the services in the service group, and the services can obtain one or more spatial streams according to their own transmission rate requirements. For example, when $MCS = 9$, the maximum transmission rate of a single-stream 106-tone RU is about 50 Mbps [16]. If the transmission rate of a service is required to be 80 Mbps, 2 spatial streams are allocated to it, so that the transmission rate can meet the requirements of the service. Figure 4 shows the distribution of spatial streams in a RU. The service type of this service group is video and image service, and one or more spatial streams are allocated to it according to the transmission rate requirements of each service.

After the spatial stream allocation scheme is determined, the AP measures the channel characteristics from each antenna to each terminal. Then, according to the channel characteristics, the AP performs precoding calculation on the data to be sent, and sends the precoding signal on each antenna. The result is that when data reaches each terminal, it contains only the data from its own terminal, eliminating the data from other terminals, and forming a directional beam to each terminal. Multiple input antennas and multiple output antennas can simultaneously transmit different data on different spatial paths, and their spatial irrelevance brings spatial isolation of services. The isolation intensity between spatial streams is related to the antenna design of WiFi6 devices. The larger the antenna spacing, the stronger the antenna isolation.

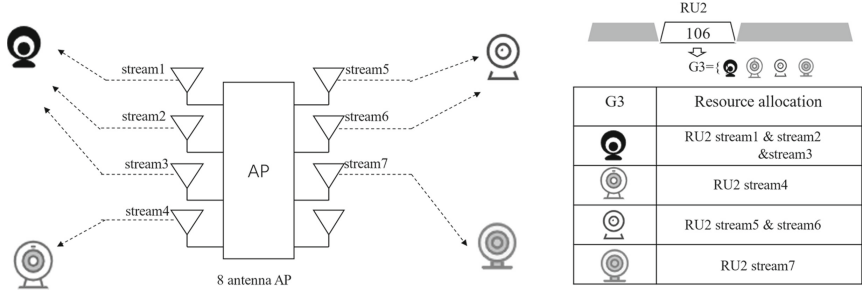


Fig. 4. MIMO spatial streams allocation method.

3.3 Isolated Resource Allocation Method Based on Improved KM Algorithm

Different types of service in the underground pipe gallery have different isolation requirements, and isolation capabilities of RU at different positions on the spectrum are also different, so the mapping between service and communication resources needs to be established according to the isolation requirements of service. In the process of resource allocation, isolation factor is introduced to test the advantages and disadvantages of allocation scheme. The isolation factor is related to the isolation and other requirements of services. In this paper, the isolation factor parameters of data collection service, video and image service, voice call service, and emergency service are respectively set to 1,5,9,25. At the same time, in order to prevent some services from obtaining communication resources for a long time, the isolation factor parameter is updated with the service waiting time, and at the end of each TXOP, the isolation factor for all services that do not receive communication resource increases by 1.

At the beginning of each round of resource allocation, the AP obtains the service information uploaded by the STA. The isolation factor of service is $L = \{L_1, L_2, \dots, L_M\}$. Divide service groups based on service types and isolation requirements, and calculate the comprehensive isolation factor $GL = \{GL_1, GL_2, \dots, GL_X\}$, corresponding to service group $G = \{G_1, G_2, \dots, G_X\}$, and GL_x is the sum of service isolation factors in G_x . Then, calculate the service guardband requirement $GV = \{GV_1, GV_2, \dots, GV_X\}$.

After the division of the service group is completed, the RU resources in the frequency domain are allocated. The quantity of 106-tone RU under the current working bandwidth is Y , and the distance between 106-tone RU and the nearest 106-tone RU on the spectrum is $DR = \{DR_1, DR_2, \dots, DR_Y\}$. At this point, the constraints on resource allocation can be abstracted into the following:

1. When $GV_x < DR_y$, G_x can get RU_y .
2. Each service group can only get one RU.
3. Each RU can only be assigned to one service group.

The objective function can be expressed as formula (4), the matching relationship between RU and service group is recorded as $GR = \{GR_1, GR_2, \dots, GR_X\}$, when G_x

gets $RU_y, GR_x = RU_y$, otherwise $GR_x = Null$.

$$\max[f(GR)] = \max \sum_{x=1}^X [GL_x * (GR_x \neq Null)] \quad (4)$$

This problem can be solved by the idea of graph theory. The problem is abstracted into a matching problem of weighted bipartite graph, and the improved Kuhn-Munkres (KM) algorithm is used to solve it. The two point sets of the bipartite graph are $G = \{G_1, G_2, \dots, G_X\}$ and $RU = \{RU_1, RU_2, \dots, RU_Y\}$. According to constraint condition ①, when $GV_x < DR_y$, there is a weighted edge $E(G_x, RU_y)$ between G_x and RU_y , and the weight is GL_x . According to the constraint condition ②③, no two edges in the final set of desired edges have the same vertices. When the sum of edge weights in the final edge set is the largest, the objective function gets the optimal solution.

1. Add “virtual edge” and “virtual point”, add “virtual point” to the point set with fewer points to make the number of centralized points of the two points the same, add “virtual edge” to transform the graph into a complete bipartite graph, the weight of “virtual edge” is 0.
2. Take a feasible vertex labelling l on the graph $Graph = (G, RU; E)$, then get the l equality subgraphs $Graph_l$ of $Graph$, which is $Graph_l$, satisfying the pair $\forall G_x \in G, RU_y \in RU$, have $l(G_x) + l(RU_y) \geq w(G_x, RU_y)$, where $w(G_x, RU_y)$ is the weight of edge $E(G_x, RU_y)$, initially set $(G_x) = \max_{y \in Y} w(G_x, RU_y)$, $l(RU_y) = 0, \forall G_x \in G, RU_y \in RU$. Any matching $Match$ (can be \emptyset) on $Graph_l$ is used as the starting match.
3. If $Match$ saturates each vertex of G , then $Match$ is the optimal match and stops; otherwise, take any vertex u that is not saturated by $Match$, let $S = \{u\}, T = \emptyset$.
4. If the neighbor set $Neighbor_{Graph_l}(S) \cap T$, go to Step 5. Otherwise, $Neighbor_{Graph_l}(S) = T$. Calculate α_l , and new feasible vertex labels \tilde{l} . Let $l = \tilde{l}, Graph_l = Graph_{\tilde{l}}$.

$$\alpha_l = \min_{G_x \in S, RU_y \notin T} \{l(G_x) + l(RU_y) - w(G_x, RU_y)\} \quad (5)$$

$$\tilde{l}(v) = \begin{cases} l(v) - \alpha_l, v \in S \\ l(v) + \alpha_l, v \in T \\ l(v), v \notin T \cup S \end{cases} \quad (6)$$

5. Select $RU_y \in Neighbor_{Graph_l}(S) \setminus T$, if RU_y is $Match$ -saturated, set $(RU_y, G_x) \in Neighbor_{Graph_l}(S)$, make $S \leftarrow S \cup \{G_x\}, T \leftarrow T \cup \{RU_y\}$, go to Step 4. Otherwise, RU_y is $Match$ -unsaturated, and there is a $Match$ -augmenting path P , let $M = M \Delta E(P)$, go to Step3.

Through the above algorithm, the $Match$ obtained is the maximum weight edge set. KM algorithm is used to solve the maximum weight problem under perfect matching [17], but the final solution result of this problem need not be perfect matching. Therefore, “virtual point” and “virtual edge” are added. As shown in Fig. 5, RU_5 is the added

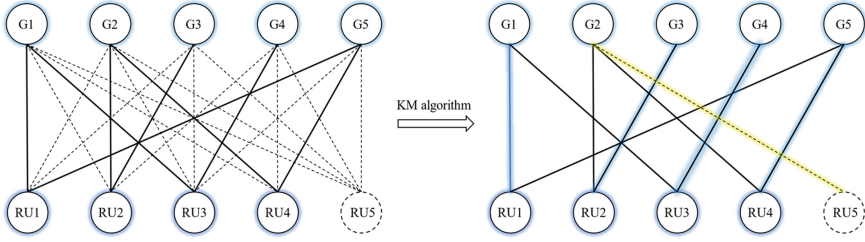


Fig. 5. Improved KM algorithm to solve RU assignment problem.

“virtual point”, and the mapping between RU5 and G2 means that G2 has not obtained RU resources.

After the allocation of RU is completed, MU-MIMO technology is used to reuse 106-tone RU to obtain several spatial streams, which are allocated according to the transmission rate requirements of each service in the service group carried by this RU. Finally, distribute the resource allocation plan to each STA. At this time, if there are services in the resource request sequence that have not been allocated resources, increase their isolation factor by one and let them participate in the next round of resource allocation.

4 Simulation

In order to verify the feasibility of the proposed multi-service isolation scheme, a simulation experiment is carried out. The test network consists of 1 AP and 30 STAs, and all devices support 8×8 MU-MIMO. The wireless channel model conforms to Rayleigh fading. Working band is 2.4 GHz, working bandwidth is 40 MHz, MCS value is 9, 256 QAM modulation, bit rate is $5/6$, Symbol bit length is 8, Symbol Time is 13.6 us (frame gap GI is 0.8 us, basic transmission Time is 12.8 us), TXOP is 5 ms.

The theoretical maximum rate of 40 MHz is $(8 * 5/6 * 468 * 8)/(12.8 + 0.8) = 1835$ Mbps, and the theoretical maximum rate of single-stream 106-tone RU is $(8 * 5/6 * 106 * 1)/(12.8 + 0.8) = 52$ Mbps. However, the actual transmission rate generally cannot reach the theoretical value, and the simulation assumes that the actual transmission rate is 50% of the theoretical transmission rate, that is, 26 Mbps.

Firstly, the performance of service isolation in frequency domain and spatial domain is verified. The isolation in frequency domain is realized by OFDMA technology and guardband. Under the working bandwidth of 40 MHz, there are four 106-tone RU. When they all transmit data, the guardband they can provide for service is 2187.5 kHz, 546.875 kHz, 546.875 kHz and 2187.5 kHz respectively. Figure 6 shows RU allocation at 40 MHz at a certain time, RU_1 and RU_4 can meet the requirements of four kinds of service isolation, RU_2 and RU_3 can meet the “Data collection” and “video and image” isolation requirements of the service. Four types of different services can find suitable RU to meet their guardband requirements.

MU-MIMO technology by making different service in different spatial path transmission to realize, so the isolation performance is associated with antenna hardware design. WiFi6 antenna isolation refers to the ratio of the signal transmitted by one antenna and

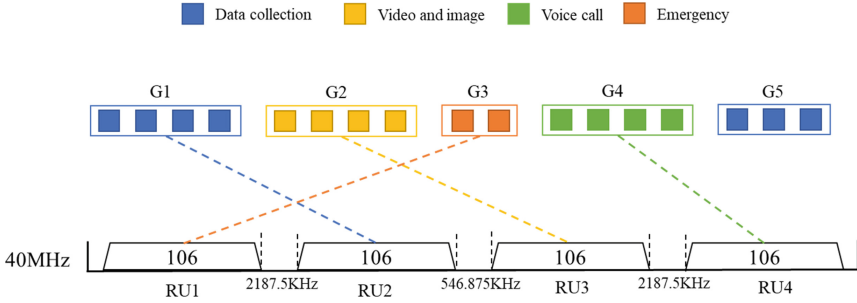


Fig. 6. The allocation of RU resources in 40 MHz.

the signal received by the other antenna to the signal of the transmitting antenna. That is, the less signal one antenna receives from the other antenna, the better the isolation between the two antennas and the lower the degree of interference. Currently, the isolation degree of WiFi6 devices on the market is generally higher than 15 dB in 2.4 GHz and 5 GHz bands, which meets the communication service isolation requirements of underground pipe gallery.

Secondly, the data transmission of different types of services is compared. Due to the large difference in the number of various communication services in the underground pipe gallery, different proportions were set for the four types of services in the simulation. Data collection services account for 50%, denoted as A in figure, video and image services account for 20%, denoted as B; Voice call services account for 20%, denoted as C, and emergency services account for 10%, denoted as D. Figure 7 shows the variation of latency with the number of STA. With the increase of the number of active STA, the average waiting time of the four types of service increases. In the vertical comparison, services with higher isolation have lower average latency. Emergency service can receive faster response from communication resources. Figure 8 shows the change of the average throughput of services with the number of STA. It can be seen that the throughput of emergency services and video and image services is larger, which is related to the higher transmission rate requirements of them.

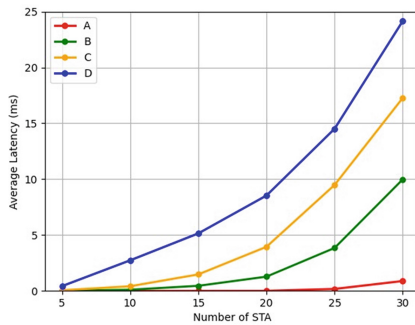


Fig. 7. Latency varies with the number of active STAs.

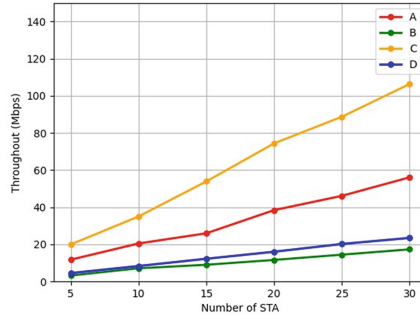


Fig. 8. Throughput varies with the number of active STAs.

Then, resource usage is compared with that without the service isolation solution. The channel division method in the comparison scheme is the same as that in the isolation scheme, but the spatial streams on 106-tone RU is not limited to the same type of service, and the resource allocation principle is FCFS (first-come-first-served). Figure 9 shows the variation of the average latency with the number of active STA of the two resource allocation schemes. The average latency of the isolation scheme is slightly higher than that of the comparison scheme, and the more active STA, the greater the latency gap. Figure 10 shows the variation of the throughput with the number of active STA of the two resource allocation schemes. When the number of active STAs is the same, the throughput difference between the two schemes is about 11%–23%, and the comparison scheme is slightly higher than the isolation scheme. Simulation results show that the resource allocation scheme has low resource cost while ensuring communication service isolation.

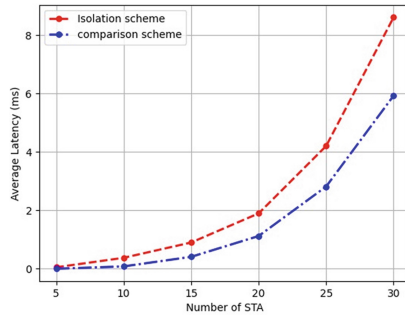


Fig. 9. Latency comparison of the two schemes.

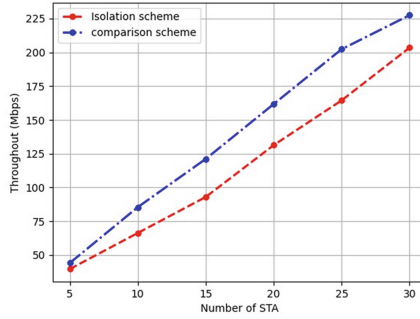


Fig. 10. Throughput comparison of the two schemes.

5 Conclusion

In this paper, a multi-service communication isolation method based on WiFi6 is designed to solve the communication isolation problems of different services in underground pipe gallery. Compared with the traditional wireless communication isolation scheme, the service isolation scheme in this paper is implemented based on the 802.11ax protocol, with low cost and strong scalability. It mainly uses the channel characteristics of WiFi6 frequency domain and spatial domain. This paper proposes a communication multi-service physical isolation architecture based on WiFi6 is proposed to realize the physical isolation of communication services from the two levels of frequency domain and spatial domain. In the isolation process, services are divided into service groups according to different service isolation requirements. In the frequency domain, the spectrum isolation scheme based on guardband is used to isolate different types of services. In the spatial domain, the isolation of same type services is realized by the spatial stream isolation scheme based on service grouping, and the isolation resource allocation mechanism is established through the improved KM algorithm based on the idea of graph theory. Through simulation verification, the solution can provide reliable isolation guarantee for services with different isolation requirements, and services with high isolation requirements such as emergency services can quickly receive resource responses, and pay a small resource cost while ensuring communication isolation.

Acknowledgements. This work is supported by State Grid Jiangsu Electric Power Co., Ltd. Science and Technology Project “Research on Low Cost Wireless Coverage and Trusted Access Technologies for Underground Pipe Gallery Digital Network” (J2021081).

References

1. Jinpeng, H., Yanping, Z., Xiaoyi, H., Liu, J., Jianwei, C.: Application prospect of electromagnetic communication in intelligent pipe gallery. In: 2019 International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM), pp. 1–2 (2019)
2. Zhuang, X., Chen, S., Wu, J., Wang, C.: Underground comprehensive pipe gallery illegal intrusion wireless sensing system. In: 2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1–5 (2018)

3. Maldonado, R., et al.: Comparing Wi-Fi 6 and 5G downlink performance for industrial IoT. *IEEE Access* **9**, 86928–86937 (2021)
4. Singh, M., Baranwal, G.: Quality of service (QoS) in internet of things. In: 2018 3rd International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), pp. 1–6. IEEE (2018)
5. Bellalta, B.: IEEE 802.11ax: high-efficiency WLANs. *IEEE Wirel. Commun.* **23**(1), 38–46 (2016)
6. Chen, Q., Zhu, Y.-H.: Scheduling channel access based on target wake time mechanism in 802.11 ax WLANs. *IEEE Trans. Wirel. Commun.* **20**(3), 1529–1543 (2020)
7. Qu, Q., Li, B., Yang, M., Yan, Z.: An OFDMA based concurrent multiuser MAC for up-coming IEEE 802.11ax. In: 2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), pp. 136–141 (2015)
8. Taranez, M., Ikuno, J.C., Rupp, M.: Sensitivity of OFDMA-based macrocellular LTE networks to femtocell deployment density and isolation. In: ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems. VDE (2013)
9. Heo, Y., Jang, J., Kim, Y., Yang, H.J.: Performance comparison of SU- and MU-MIMO in 802.11ax: delay and throughput. In: 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 879–882 (2020)
10. Jacobsson, S., et al.: Massive MU-MIMO-OFDM uplink with hardware impairments: modeling and analysis. In: 2018 52nd Asilomar Conference on Signals, Systems, and Computers. IEEE (2018)
11. Khorov, E., Kiryanov, A., Lyakhov, A., Bianchi, G.: A tutorial on IEEE 802.11ax high efficiency WLANs. *IEEE Commun. Surv. Tutor.* **21**(1), 197–216 (2019, firstquarter)
12. Kulkarni, J., Desai, A., Sim, C.Y.-D.: Wideband four-port MIMO antenna array with high isolation for future wireless systems. *AEU-Int. J. Electron. Commun.* **128**, 153507 (2021)
13. IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks–Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Enhancements for High-Efficiency WLAN. In: IEEE Std 802.11ax-2021 (Amendment 1: Enhancements for High-Efficiency WLAN), pp. 1–767, 19 May 2021 (2021). <https://doi.org/10.1109/IEEESTD.2021.9442429>
14. Oughton, E.J., et al.: Revisiting wireless internet connectivity: 5G vs Wi-Fi 6. *Telecommun. Policy* **45**(5), 102127 (2021)
15. 3GPP TR 38.802 v. 14.2.0 Study on New Radio (NR) Access Technology; Physical Layer Aspects. Technical Specification Group Radio Access Network Release 14, September 2017
16. Masiukiewicz, A.: Throughput comparison between the new HEW 802.11 ax standard and 802.11 n/ac standards in selected distance windows. *Int. J. Electron. Telecommun.* **65**(1), 79–84 (2019)
17. Cui, H., Zhang, J., Cui, C., Chen, Q.: Solving large-scale assignment problems by Kuhn-Munkres algorithm. In: Proceedings of 2nd International Conference on Advanced Research in Mechanical Engineering and Industrial Informatics, pp. 822–827 (2016)



MU-PDR: A Method of Fingerprint Passive Positioning for WiFi6 Based on MU-RTS/CTS

Ningchi Zhang^{1,3}, Cheng Zhong², Qinghai Ou^{1,3}, Pengcheng Lu², La Bao³, Yan Liu⁴,
and Sujie Shao⁴(✉)

¹ Beijing Fibrlink Communications Co., Ltd., Beijing 100070, China

² Xiong'an New District Power Supply Company of State Grid Hebei Electric Power Co., Ltd.,
Xiong'an 071600, Hebei, China

³ State Grid Information and Communication Industry Group Co., Ltd., Beijing 102211, China

⁴ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts
and Telecommunications, Beijing 100876, China
buptssj@bupt.edu.cn

Abstract. With the rapid development of wireless communication technology and mobile Internet, requirements for accuracy and efficiency of location services are on the increase. In recent years, WiFi fingerprint passive positioning technology has received more and more attention from the academic community. RSS fingerprint is usually adopted in this field, but it has some inherent defects in real-time positioning due to its unstable acquisition timing. Packet Delivery Rate (PDR) fingerprint has been proposed in recent years, which has great advantages over RSS fingerprint in passive positioning scenario. However, it still has the problems of low positioning efficiency and huge resource consumption. Therefore, this paper proposes an improved positioning method named MU-PDR (Multi User Packet Delivery Rate), which utilizes the newly introduced MU-RTS/CTS feature in WiFi6. The corresponding MU-PDR fingerprint acquisition and positioning methods are proposed to optimize the performance of PDR-like fingerprint in multi-target scenes. Finally, simulation result shows that MU-PDR has higher efficiency than traditional PDR in multi-target scenario, which is illustrated by the significant reduction of channel resource occupation under the same accuracy requirements.

Keywords: Fingerprint positioning · Passive positioning · WiFi6

1 Introduction

With the popularity of WiFi devices, WiFi based positioning technology [1] shows significant advantages of low cost and easy deployment, which has excellent development potential. WiFi-based fingerprint positioning technology realizes the positioning and tracking of user terminals through different signal characteristics and matching algorithms [2]. This technology usually does not need additional hardware support and can be completely based on common WiFi devices. It is commonly realized through received

signal strength (RSS) [3]. The advantage of location fingerprint technology is that it can handle the difference of signal characteristics in different locations without signal resolution. It is not affected by non-line-of-sight propagation (NLOS) errors and has excellent stability and robustness.

According to which side the signal is acquired and the fingerprint is generated, the fingerprint positioning can be divided into active method and passive method [4]. Active positioning covers most of the current fingerprint positioning technologies [5], that is, collecting data sent by AP on the positioning target side, generating location fingerprints and matching the fingerprints of the server-side fingerprint database. In contrast, passive fingerprint technology collects the wireless signal sent by the target on the AP side, generates a fingerprint and uploads it to the local server for matching and estimating the location of the signal source. Its characteristic is that the AP side has the initiative of positioning, and the user usually does not need to install software and hardware or proceed with something else on the mobile device. The positioning process is semi-independent from the user's participation.

At present, passive fingerprint positioning [6] mostly adopts RSS as the fingerprint, and mainly acquires RSS data through probe request frame and data frame. Due to the long transmission interval of the probe request frame, the frequency of RSS obtained by the positioning system is low, and vacuum period is prone to occur. Moreover, the data frame also has problems such as unstable sending frequency and loss of key information. The AP must capture the data frame sent by the target to parse the RSS data contained in it. In practice, it is rare for the target to actively send packets, and the frequency with which the target sends packets is unpredictable, especially when the target is not associated with any AP. In this case, the positioning system can not get RSS of a specific device at any time, and the real-time requirement can not be satisfied.

For such problems, Duan et al. [7] proposed a positioning technology solution based on the Packet Delivery Rate (PDR), where the AP actively sends a request to send (RTS) frame to solicit the target device to reply to a clear to send clear to send (CTS) frame to calculate the PDR, and use this instead of RSS as the fingerprint. While satisfying the real-time requirement of passive positioning, they solve the device dependence problem by specifying modulation mode and power. Since the PDR positioning technology adopts the RTS/CTS mechanism, a group of APs in the positioning system can only locate a single target in a same period. The high-frequency and mass-transmitted positioning frames will inevitably lead to waste of channel resources in the multi-target scenario. As a result, the cost in real scenarios is worth paying attention to. In general, under the previous protocol, the one-to-one mechanism of RTS/CTS is the most important bottleneck limiting this positioning technology.

In this paper, we propose a passive positioning method of MU-PDR (Multi User Packet Delivery Rate) fingerprint integrated with the MU-RTS/CTS [8] features newly introduced in 802.11ax [9, 10], based on the passive positioning technology of PDR fingerprints. This method integrates the more flexible channel resource allocation mechanism in 802.11ax into the PDR fingerprint positioning technology, and replaces the traditional RTS with MU-RTS, so that the target can perform parallel CTS responses, which greatly improves the condition of the same transmission data rate. It improves the efficiency of fingerprint collection and construction, thus expanding the potential

of PDR fingerprint positioning technology in multi-target scenarios. Through simulation experiments, this paper verifies the effectiveness of the method from two aspects: positioning accuracy and the number of positioning frames. The results show that the positioning accuracy of MU-PDR method for multiple targets is basically the same as that of traditional PDR method for single target, but the overall positioning cost of the former is significantly reduced.

2 MU-PDR Fingerprint Positioning Frame

The MU-PDR fingerprint acquisition method is based on the MU-RTS/CTS mechanism in WiFi6. The parallel PDR data calculation is realized through the CTS replied in parallel in multiple RU. The allocation mode of positioning RU adapts to the general 802.11ax protocol. MU-PDR positioning method is a multi-AP collaborative positioning process. Therefore, it is necessary to plan the transmission time of positioning frames reasonably, in order to avoid the waste of resources and accuracy loss caused by the collision of high-frequency positioning frames from different APs (Fig. 1).

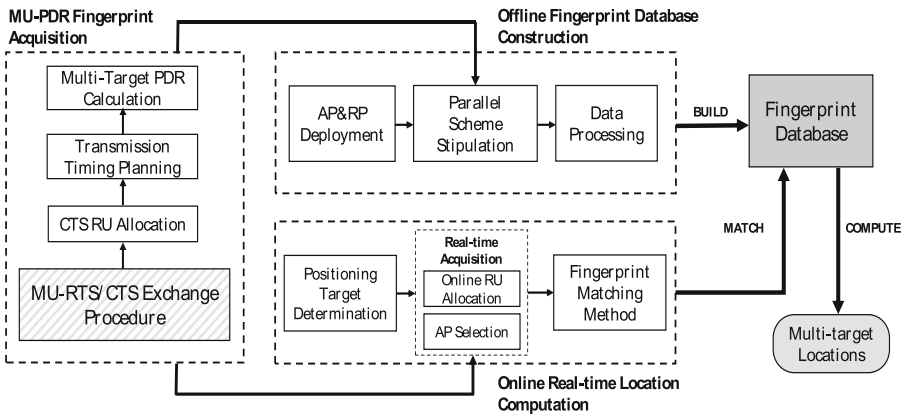


Fig. 1. The PDR involved in the online and offline phase is obtained by calculating the MU-RTS frame sent by the AP and the CTS frame replied by multiple targets in parallel.

The construction method of offline MU-PDR fingerprint database is similar to the general one of fingerprint database. The positioning range is planned according to the AP coverage, and some location points in the indoor space are preselected as reference points (RP). Based on the MU-PDR fingerprint acquisition method, the multi-target PDR fingerprint of an RP is collected in parallel to form the complete MU-PDR fingerprint as the location feature of the RP. After data processing, the MU-PDR fingerprints data of multiple APs are stored in the local server to build an offline fingerprint database.

In the section of online real-time location computation, it is first necessary to collect real-time MU-PDR fingerprints, in a multi-AP and multi-target scenario, which involves the determination of targets before positioning, the positioning RU planning of each target, and the specific enabled APs in this positioning, so as to ensure the positioning

efficiency and reduce the occupation of channel resources. Finally, the MU-PDR fingerprint collected in real time is matched with the fingerprint database to obtain the positioning results. Different matching algorithms have some impact on the accuracy.

3 MU-PDR Fingerprint Acquisition Method

The behavior of different types of WiFi terminals is uncertain. Thus, it is difficult to ensure the accuracy and timeliness of information acquisition by using the probe request frames and data frames from the terminals. One way to solve this problem is to install corresponding programs on the WiFi terminal device to cooperate with fingerprint collection on the side of the positioning system, but this way obviously violates the original intention of passive positioning. Therefore, a more reasonable way is to make the AP solicit the terminal to a certain extent within the scope of the protocol.

3.1 MU-RTS/CTS Mechanism

MU-RTS is a new trigger frame in the 802.11ax. The trigger frame requests and allocates resources for one or more HE-TB-PPDU (High Efficient Trigger-based Physical Layer Protocol Data Unit) transmissions, and carries other information required by the responding STA to send the HE-TB PPDU.

802.11ax APs firstly need to initiate a TXOP transmission time through competition (CSMA/CA). During this TXOP period, the channel is occupied by this AP. Then, this AP reserves the channel by sending MU-RTS to terminals and avoids some issues such as “hidden terminal”.

When receiving an MU-RTS, the terminal should feed back the CTS to the AP for confirmation. The MU-RTS frame contains a list of RU assignments for each 802.11ax client and helps coordinate the multiuser frame exchange. The 802.11ax clients send clear-to-send (CTS) responses in parallel using their assigned RUs, thereby enabling parallel transmission of CTS frames from multiple terminals. By setting the NAV timers of all other nodes, all terminals are in a passive receiving state and will not actively compete for channels within the reservation time. The time value set by the NAV timer is used for OFDMA data frame exchange procedure (Fig. 2).

Compared with the traditional RTS/CTS mechanism, MU-RTS/CTS mechanism provides more flexible resource allocation and channel reservation ability, which logically changes from one-to-one to one-to-many. It can greatly improve the measurement efficiency of PDR in multi-target scenario.

3.2 MU-PDR Acquisition Method

MU-PDR fingerprints essentially differentiate locations by differences in the success rate of MU-RTS/CTS exchanges at different physical locations. PDRs from multiple APs are combined to form MU-PDR fingerprints with positioning capabilities. Furthermore, the APs need to specify lower CTS reception power in the corresponding field of MU-RTS to ensure the location discrimination and the equal CTS transmission power between different terminals.

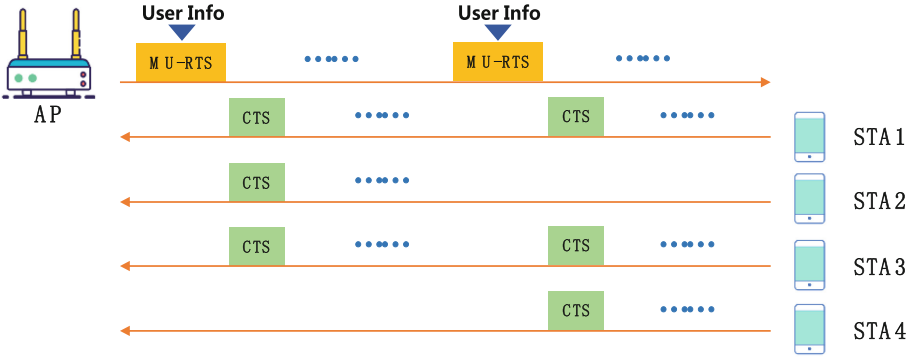


Fig. 2. MU-RTS transmits on the entire channel, which can be received by both 802.11ax nodes and traditional 802.11 nodes.

The process of a single MU-PDR fingerprint collection is that the AP sends the MU-RTS frame to the target at the specified sampling rate in the specified time window, and receives the parallel CTS of each target.

Define the sampling window size T as the time used for one fingerprint collection, and the data rate s as the MU-RTS sampling rate of a single AP within the window period, then the number of MU-RTS frames sent by a single AP for a single acquisition of a single target can be calculated as:

$$r = T \times s \tag{1}$$

Assuming that the number of CTS frames replied by this target received by the AP after the end of the sampling window is c ($c \leq r$), the PDR can be calculated as:

$$PDR = \frac{c}{r} \tag{2}$$

Consider a fingerprint collection of m targets by n APs. After the window period ends, the PDR of each target is calculated separately and combined as:

$$\begin{bmatrix} PDR_1^1 & PDR_1^2 & \dots & PDR_1^n \\ PDR_2^1 & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ PDR_m^1 & \dots & \dots & PDR_m^n \end{bmatrix} \tag{3}$$

where PDR_j^i is the PDR collected from the AP_i to the $Target_j$. If m is greater than the maximum number of parallels in the scheme, it is ensured that each target has the same positioning resources by increasing the data rate.

Since MU-PDR fingerprint collection requires high-frequency MU-RTS/CTS exchange procedure. In a positioning process with dense APs, it is necessary to reduce the interference between positioning frames sent by different APs as much as possible, which will occupy time resources and even have some impact on the positioning accuracy. When used for positioning, a MU-RTS/CTS exchange procedure does not initiate

an uplink transmission of the terminal at the end, but defaults to relinquishing the channel and competing for the next transmission opportunity.

Next, Consider the time consumed for a single MU-RTS/CTS exchange:

- a. If the AP sends MU-RTS and successfully receives CTS frames addressed by MU-RTS trigger frame:

$$T_{success} = MU_RTS_TIME + CTS_TIME + aSIFSTime \quad (4)$$

- b. The 802.11ax protocol states that the AP should wait for a CTSTimeout interval after sending MU-RTS:

$$CTSTimeout = aSIFSTime + aSlotTime + aRxPHYStartDelay \quad (5)$$

The interval starts with the PHYTXEND.confirm primitive of the MU-RTS trigger frame sent by the MAC. If the MAC does not receive the PHY-RXSTART.indication primitive during the CTSTimeout interval, the AP should determine that the transmission of MU-RTS trigger frames failed.

Therefore, this paper stipulates that MU-RTS should set a very low NAV(TXOP_DURATION) for other APs and terminals in the same channel, just enough to protect this exchange.

In the MU-PDR positioning method proposed in this paper, due to the high frequency of positioning frames, the failed exchange caused by other positioning frames at the same time should be avoided, and the PDR obtained in the online phase should not be affected by the number of enabled APS.

This method preliminarily optimizes this process by separating the MU-RTS transmission timing of each AP. Assuming a fingerprint acquisition initiated by N APs, the transmission window is divided into $B \times N$ blocks. Each block is allocated to each AP in order, and each AP only performs MU-RTS/CTS procedure for positioning purpose in the allocated time block (Fig. 3).



Fig. 3. In this way, the acquisition process of MU-PDR fingerprint becomes more fine-grained, and it is possible to further optimize according to the PDR data of each reduced window.

Then the specific sampling window of AP_i is defined as:

$$\left[\frac{bn+i-1}{Bn} \cdot T, \frac{bn+i}{Bn} \cdot T \right], \quad b = 0, \dots, B-1 \quad (6)$$

The sampling rate in a single reduced time window is $n \cdot s$.

By splitting multiple blocks, the MU-RTS/CTS exchange of each AP is independent of each other, and the independent transmission window is evenly distributed in each part of the total window. The simulation part of this paper will verify the improvement of this scheme in time domain resource occupation.

The above MU-PDR fingerprint construction method is carried out within the scope of 802.11ax protocol, without pre-operation of terminal equipment, and has good applicability to 802.11ax terminals in general scenes, with low deployment difficulty.

4 MU-PDR Fingerprint Positioning Method

4.1 Offline Fingerprint Database Construction Method

In the offline phase, the indoor area should be divided firstly. Based on the actual situation, the area is divided into RPs separated by a certain distance. The granularity and distinction of RPs need to be considered simultaneously. Usually the distance between adjacent RPs is about 1 m, and the final positioning results are indirectly from the locations of these RPs.

For convenience, the following discussion defaults to a 20 MHz channel. Four bandwidth sizes of RUs are defined in 20 MHz channels: 2 MHz (RU_a), 4 MHz (RU_b), 8 MHz (RU_c), and 20 MHz (RU_d). The impact of different RU is reflected not only in throughput, but also in packet delivery rate. Therefore, RUs with different bandwidth sizes should be attributed to different fingerprint dimensions and have corresponding location discrimination capabilities.

Given a positioning system consisting of N AP. Within the coverage of this positioning system, for each RP point, the AP measures the combined PDR of the point by multiple MU-RTS/CTS exchanges, as shown in Eq. (3). In this phase, the set of targets is in the same location, thus obtaining PDR data of each size of RU from an AP at that RP.

$$PDR_{(AP_i, RP_j)} = [PDR_{RU_a}, PDR_{RU_b}, PDR_{RU_c}, PDR_{RU_d}] \quad (7)$$

Then combine all AP data to get the MU-PDR fingerprint of the RP.

$$MU - PDR_{RP_j}^{Offline} = \begin{bmatrix} PDR_{(AP_1, RP_j)} \\ PDR_{(AP_2, RP_j)} \\ \dots \\ PDR_{(AP_N, RP_j)} \end{bmatrix} \quad (8)$$

For this RP_j , PDR_x^y represents the PDR measurement value of the x_{th} AP in RU_y . The MU-PDR fingerprint serves as the complete location feature for this RP_j .

The MU-PDR fingerprints and their coordinates for all RPs construct a fingerprint database that is stored on the local server for invocation during the online phase. Considering the complexity of the indoor environment, a more representative MU-PDR fingerprint is usually measured multiple times at each RP.

4.2 Online Real-Time Location Computation Method

Online Fingerprint Obtaining and Matching

In the online phase, a scene with multiple positioning targets is given. After determining multiple targets to locate by positioning requests from the upper application, this positioning system will perform the RU pre-allocation process of OFDMA according to the MU-PDR acquisition method described above, assigning a specific band to each locating target (Fig. 4).

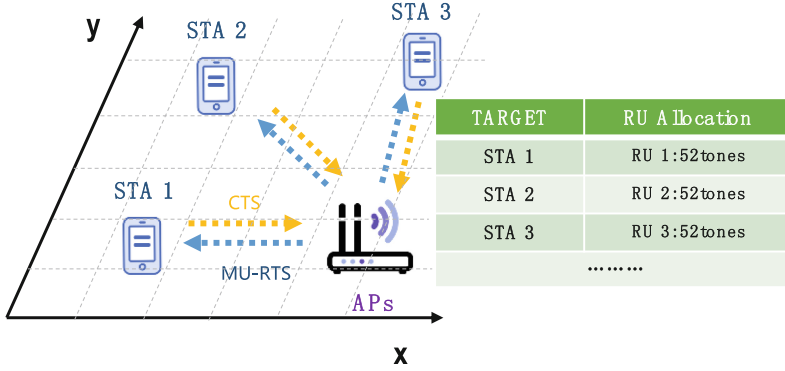


Fig. 4. An exchange process involving three STAs at a time.

Next, the system selects the appropriate sampling window and sampling rate based on the actual situation. Within a certain range, the higher the MU-RTS sampling rate is, the more stable the PDR data will be, and the more channel resources will be occupied.

Consider a single localization initiated by n ($n \leq N$) APs to m targets. When the sampling window is over, MU-PDR fingerprints are calculated based on the total number of MU-CTS sent and the number of CTS received by each allocation channel:

$$MU - PDR^{Online} = \begin{bmatrix} PDR_{AP_1}^{Online} \\ PDR_{AP_2}^{Online} \\ \dots \\ PDR_{AP_n}^{Online} \end{bmatrix} \quad (9)$$

$$PDR_{AP_i}^{Online} = [pdr_{Tgt_1}^{RU_1}, pdr_{Tgt_2}^{RU_2}, \dots, pdr_{Tgt_m}^{RU_m}] \quad (10)$$

where pdr_{Tgt_j} represents the PDR of the target Tgt_j over an allocated RU by AP_i .

Then, using the computed MU-PDR fingerprint and the MU-PDR fingerprint database constructed in the offline phase, the positioning system performs a matching algorithm to calculate the results. The specific matching algorithm will have a slight impact on the positioning accuracy. As an example, the KNN algorithm is adopted in this paper, by selecting the K most similar locations and average their coordinates (Fig. 5).

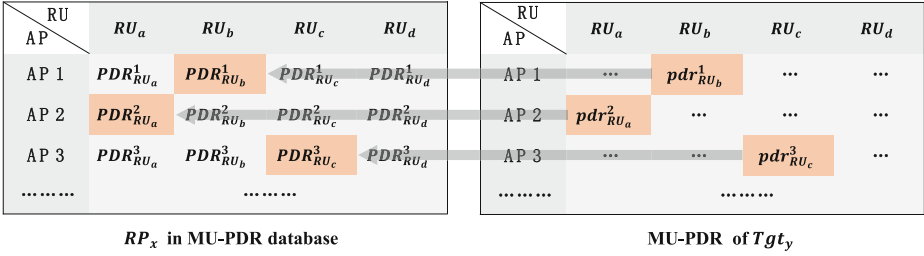


Fig. 5. The collected online fingerprint and the offline fingerprint at the corresponding RP.

About MU-PDR fingerprints involved in this phase, the column vectors represent each target’s PDR fingerprint, and each dimension of the fingerprint can find the corresponding in the offline fingerprint database. Similarity is calculated from the Euclidean distance between the offline and online fingerprints:

$$SIM(RP_x, Tgt_y) = \frac{1}{\sqrt{\sum_{i=1}^n (PDR^i_{RP_x} - pdr^i_{(Tgt_y, RU)})^2} + \epsilon} \tag{11}$$

Select the K RPs with the greatest similarity of the positioning target, and compute the location estimates of multiple targets at the same time. Therefore, the multi-target parallel positioning initiated by the positioning system is realized.

$$\begin{cases} (x_1, y_1) = \frac{\sum_{i=1}^K (x_1^i, y_1^i)}{K} \\ \dots \dots \\ (x_m, y_m) = \frac{\sum_{i=1}^K (x_m^i, y_m^i)}{K} \end{cases} \tag{12}$$

Quantity Limit of Concurrent Enabled AP

The accuracy of MU-PDR fingerprints depends on the number of MU-RTS/CTS sent. However, a higher amount of transmission may result in higher channel time-frequency resource consumption and power consumption.

Set p as the number of CTS parallels on the target side and $p = 1$ as the case with traditional PDR schemes. It is easy to get the increase of time-frequency resource occupancy of MU-PDR scheme compared with traditional PDR scheme. For the same number of positioning frame exchanges, the MU-PDR scheme achieves p times the number of positioning targets compared with the traditional PDR scheme, while the time-frequency resource consumption is basically the same.

Furthermore, this paper analyses the power consumption from the point of view of the number of positioning frames. Considering the two parts of MU-RTS sent by AP and CTS sent by target, the total theoretical transmission of positioning frames in a sampling window is calculated as:

$$Frames = \underbrace{n \times T \times s \times \left\lceil \frac{m}{p} \right\rceil}_{MU-RTS} + \underbrace{n \times T \times s \times m}_{CTS} \tag{13}$$

where p is the maximum number of CTS parallel scheme. $p = 1$ represents the case with traditional PDR scheme.

$$Frames_{PDR} = 2 \times n \times T \times s \times m \quad (14)$$

Compared with traditional PDR, MU-PDR first significantly reduces the number of RTS frames sent for positioning in multi-target scenarios with the same sampling rate. It is easy to calculate that under the same other conditions, MU-PDR theoretically reduces the number of positioning frames by 25%–50%, depending on the specific parallel scheme.

When the target is within the scope of multiple APs ($N > 3$), the number of used AP needs to be limited. Although more AP means higher fingerprint dimensions, positioning accuracy can be slightly improved (with severe marginal effects). However, in dense scene of MU-PDR fingerprint positioning, it is easy to overflow the positioning frames in the channel, causing serious congestion.

To further optimize the cost of positioning frames, the number of AP used for fingerprint collection is limited. In an independent N AP positioning system, the number of AP per positioning is limited as:

$$AP_{adopt} = \min \left(AP_{cover}, AP_{max}, \frac{N}{L+Q-1+\frac{N}{(N-3)\sqrt{p}}} \frac{1}{\sqrt{p}} + AP_{min} \right) \quad (15)$$

where AP_{cover} is the number of AP around this target, AP_{max} is the maximum number of AP allowed, L is the number of targets already in the window, and Q is the number of targets of this positioning. When there are fewer targets in the positioning window, more APs are enabled for this positioning task. As the number of targets in the positioning window increases, the number of concurrently enabled AP quickly converges to the minimum fingerprint dimension requirement number AP_{min} .

MU-PDR, as a fingerprint positioning method, also has the advantages of strong robustness and no need of signal computation. MU-PDR realizes real-time passive positioning through MU-RTS/CTS exchange procedures initiated on the side of the positioning system, independent of random data frames sent by the terminal device.

5 Simulation

In this chapter, the validity of the proposed method is verified through simulation experiments, based on MATLAB. The simulation mainly includes positioning accuracy and positioning cost.

The indoor WiFi signal propagation scene is built based on MATLAB simulation platform, in which the simulation network with multiple APs and nodes is built, and the exchange process of MU-RTS/CTS mechanism is simulated. The indoor propagation model is used to determine the path loss between nodes, with the parameters of indoor scenario and the distance between nodes as the main parameters. To simulate the real indoor environment, this simulation experiment adds random ambient noise interference to the signal, and simulates the fluctuation of indoor positioning while guaranteeing the distance difference.

5.1 Positioning Accuracy

This section aims to test the impact of one-to-many mechanism of MU-RTS/CTS on positioning accuracy in multi-target scenarios.

Set a 10 m * 10 m * 3 m indoor area and divide the X/Y plane into 100 reference points, 1 m apart. The positioning system is built with three APs, which are distributed indoors and triangulated to ensure unambiguous positioning results. All APs and targets separately use the constant transmission power during this simulation.

Three schemes are compared with the same sampling window of 1 s and sampling window of 500 per second:

- Traditional RTS/CTS scheme. Locate only one single target at a time in the whole 20 MHz channel.
- Traditional RTS/CTS scheme. Considering multi-target scenario with 4 targets, the window size is compressed to 1/4 of the original.
- The MU-PDR scheme proposed in this paper, which uses 20 MHz channel for MU-RTS and 4 MHz RUs for each CTS.

The experiment simulates the offline and online phases and uses the following MU-PDR fingerprint acquisition methods: the positioning system sends MU-RTS or RTS frames to indoor targets to be located, then the targets reply CTS frames in the indicated RUs when the reception is successful. The channel environment does not change during one exchange procedure. Repeat this process in the sampling window at a given sampling rate.

During the positioning process, the target to be located is set to move with a random track in the indoor space with a constant height. After repeated positioning for many times, the average error is calculated by measuring and calculating the positioning results of each time. Finally, the positioning accuracy of each scheme under given conditions is obtained (Fig. 6).

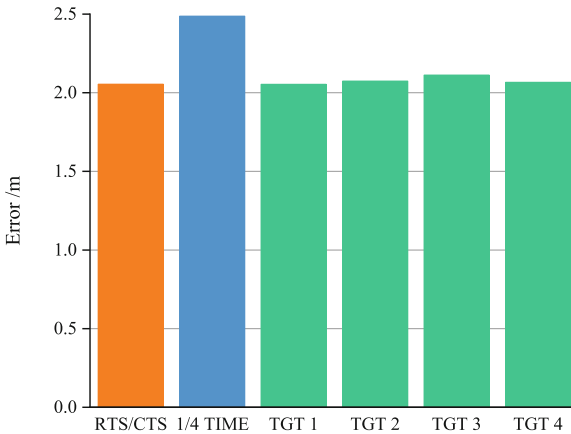


Fig. 6. Positioning accuracy simulation results.

As shown above, taking 4 MHz RU as an example, the accuracy of multi-target positioning using MU-RTS/RTS is basically the same as that of single-target RTS/CTS in this simulation environment. The results show that the CTS frames transmitted in parallel have no significant impact on the positioning accuracy. In the multi-target scenario, when the sampling window of the RTS/CTS scheme is shortened, the average error increases significantly. It can be inferred that in such PDR-like fingerprint positioning scheme, the discrimination of position is less sensitive to the transmission bandwidth of CTS than to the sampling window and the sampling rate. With similar positioning accuracy, more attention should be paid to the resources occupied in the positioning process.

5.2 Positioning Cost

This section considers a scenario where there are more APs and more targets on such region and simulates the positioning cost at the system level. Each target to be located randomly sends out a positioning request at some moments with a certain probability, and the location service is provided by some of these APs in the system. Set the same sampling window and sampling rate as the previous experiment, to simulate the number of positioning frames needed and the average time consumption of AP under different preset parameters. In order to more intuitively reflect the channel resource occupation, it is assumed that all nodes are in the same channel (Table 1 and Fig. 7).

Table 1. Simulation parameters

Parameter	Value
Sampling window	1 s
Sampling rate	1000 frames/s
CTS parallels	4
Request rate	0.1
Block num	5
$T_{success}$	25 μ s
$T_{timeout}$	59 μ s

The positioning power consumption is weighed from the number of positioning frames sent. PDR and MU-PDR use the same AP selection scheme. The simulation results display the sum of RTS (MU-RTS) and CTS frames under ideal conditions, reflecting the combined costs of different positioning schemes from the side. The result shows that when the number of targets in the area is small, there is little difference between the number of theoretical frames used for PDR positioning and MU-PDR positioning. In the case of more dense targets, the parallel construction mechanism of MU-PDR has obvious advantages in power consumption, reducing the number of positioning frames by 33.7% (Fig. 8).

In terms of the time-frequency resource consumption, this experiment tests different strategies. Different numbers of APs compete to occupy channels during a given period.

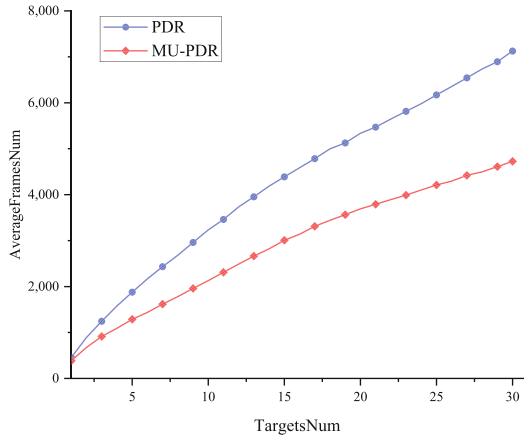


Fig. 7. Number of positioning frames simulation results.

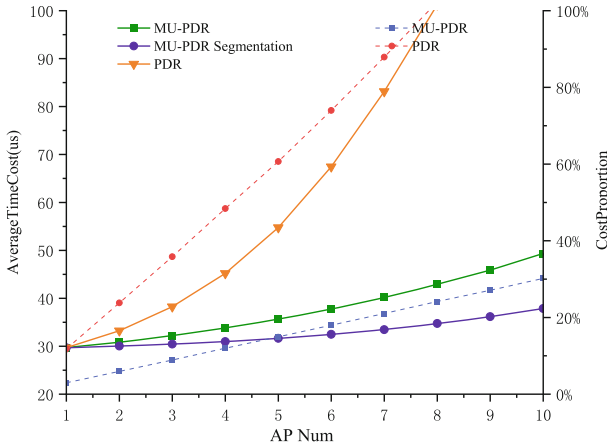


Fig. 8. Positioning time cost simulation results.

The period from waiting for transmission to exchange completion or timeout is recorded as the time consumed for one positioning procedure. In the traditional PDR scheme, the transmission success rate is low and the queuing time is long due to the excessive number of positioning frames sent and the overcrowded channel. As the number of positioning AP adopted increases, the average time consumption of a single RTS/CTS exchange will exceed 100 μ s. In this case, the time-frequency resources are insufficient to meet the sampling rate requirements. By contrast, the MU-PDR scheme improves significantly in this field. The time consumption of a single exchange is limited to less than 50 μ s. And in extreme cases, the channel occupancy is less than 30%, which is of more practical worth.

6 Conclusion

In this paper, we propose a MU-PDR fingerprint passive positioning method, which utilizes the new feature of MU-RTS/CTS in 802.11ax to increase the number of locatable targets in a single window. Compared with traditional PDR method, the positioning efficiency in multi-target scenes is improved while the positioning accuracy is guaranteed. The MU-PDR method is applicable to the general 802.11ax protocol, which does not require additional hardware support, and has the advantages of low cost and easy deployment.

Through simulation experiments, this paper verifies the validity of the scheme at the system level. The results show that the MU-PDR method proposed in this paper achieves simultaneous multi-target parallel positioning without sacrificing accuracy, and optimizes the limitation of traditional PDR fingerprint channel resource consumption, thereby greatly reducing the cost of positioning in multi-target scenes.

Acknowledgments. This work is supported by the Science and Technology Project of State Grid Corporation of China: Research and application of key technologies for wireless communication network coverage of urban power underground pipe gallery (Grant No. 5700-202113189A-0-0-00).

References

1. Basri, C., El Khadimi, A.: Survey on indoor localization system and recent advances of WiFi fingerprinting technique. In: 2016 5th International Conference on Multimedia Computing and Systems (ICMCS), pp. 253–259 (2016)
2. Chapre, Y., Ignjatovic, A., Seneviratne, A., Jha, S.: CSI-MIMO: indoor WiFi fingerprinting system. In: 39th Annual IEEE Conference on Local Computer Networks, pp. 202–209 (2014)
3. Tian, X., Shen, R., Liu, D., Wen, Y., Wang, X.: Performance analysis of RSS fingerprinting based indoor localization. *IEEE Trans. Mob. Comput.* **16**(10), 2847–2861 (2017)
4. Kim, J., Han, D.: Passive WiFi fingerprinting method. In: 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8 (2018)
5. Bahl, P., Padmanabhan, V.N.: RADAR: an in-building RF-based user location and tracking system. In: Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), vol. 2, pp. 775–784 (2000)
6. Ko, D., Kim, M., Son, K., Han, D.: Passive fingerprinting reinforced by active radiomap for WLAN indoor positioning system. *IEEE Sens. J.* **22**(6), 5238–5247 (2022)
7. Duan, Y., Lam, K.-Y., Lee, V.C.S., Nie, W., Li, H., Ng, J.K.-Y.: Packet delivery ratio fingerprinting: toward device-invariant passive indoor localization. *IEEE Internet Things J.* **7**(4), 2877–2889 (2020)
8. Khorov, E., Kiryanov, A., Lyakhov, A., Bianchi, G.: A tutorial on IEEE 802.11ax high efficiency WLANs. *IEEE Commun. Surv. Tutor.* **21**(1), 197–216 (2019)
9. Storrer, L., et al.: Indoor tracking of multiple individuals with an 802.11ax WiFi-based multi-antenna passive radar. *IEEE Sens. J.* **21**(18), 20462–20474 (2021)
10. Machrouh, Z., Najid, A.: High efficiency WLANs IEEE 802.11ax performance evaluation. In: 2018 International Conference on Control, Automation and Diagnosis (ICCAD), pp. 1–5 (2018)



The Optimal Layer of User-Specific Reconfigurable Intelligent Surfaces Structure for Uplink Communication System

Zihao Wen¹, Zhicheng Dong^{1(✉)}, Weixi Zhou², and Jing Zhao³

¹ School of Information Science and Technology, Tibet University, Lhasa 850000, China
dongzc666@163.com

² School of Computer Science, Sichuan Normal University, Chengdu, China
zhouweixi@sicnu.edu.cn

³ School of Aerospace, Xihua University, Chengdu 610039, China

Abstract. Reconfigurable intelligent surfaces (RIS) can be utilized for enhancing the communication quality, and are regarded as a promising six-generation (6G) technique. Especially, the RIS generally is deployed as a reflector or an intelligent transmission terminal. Most of the existing works have focused on the reflective RIS. In contrast, the recently emerged user-specific RIS (US-RIS) can improve performance for users at a low cost. Unlike reflective RIS, US-RIS can control the phase of the user uplink signal when the signal penetrates US-RIS. The purpose of this paper is to investigate the energy efficient of multi-layer US-RIS-aided uplink communication system with quality of service to find the optimal layer of US-RIS. To this end, we formula a joint optimization problem over several transmit beamforming vectors. An optimal energy efficient design which is based on alternating update and successive convex approximation (SCA) is proposed. Simulation results reveal that the number of elements has an impact on energy efficiency, which affects the optimal layer.

Keywords: RIS · US-RIS · SCA · Alternating optimization

This work was supported in part by the National Science Foundation of China under Grant 61561046, in part by Central government supports the reform and development of local universities of Tibet University in 2020, in part by Special fund for the development of local universities supported by the central finance of Tibet University in 2019, in part by Tibet University Special Fund Project for the Reform and Development of Local Colleges and Universities Supported by the Central Government in 2021 (The First-Class Discipline Cultivation Project), in part by Key Research & Development and Transformation Plan of Science and Technology Program for Tibet Autonomous Region (No.XZ201901-GB-16).

1 Introduction

The Fifth-generation (5G) communication is expected to reach its limits in 2030 [1–3] due to the never-ending growth of global mobile data traffic. Although the research and development are still in the early stage, many new mobile communication technologies will appear in the future six-generation (6G) [4, 5].

Reconfigurable intelligent surface (RIS) is a revolutionary technology in the field of wireless communication and one of the most important technologies in the future 6G related technologies [6, 7].

Recently the emergence of metamaterials has spawned cutting-edge technology, RIS [8]. A planar array of some passive elements can form a RIS, where the input signal is imparted at the desired phase angle by each element independently [9]. To adapt to the dynamically changing wireless propagation environment, the reflected (or transmitted) signal can be reconfigured to propagate in its desired direction by properly adjusting the phase shifts of all elements, and the reflection factor (or loss factor) of each element can be reconfigured [10]. The advantages of RISs have been demonstrated in these actual wireless communication models [11]. Compared to traditional networks without RIS, RIS-aided networks can improve energy efficiency and capacity [12, 13]. According to the different ways of changing the propagation direction of the input signal [14], RIS can be roughly divided into two types: reflective type [15, 16] and transmissive type [17]. Precisely, the reflective type RIS reflects the input signal in the desired direction. The signal can also penetrate through the transmissive type RIS, and by changing the phase, the signal is transmitted in the desired direction. The concept of user-specific reconfigurable intelligent surfaces (US-RIS) is first proposed by [18]. In the communication system, RIS is traditionally used between the base station (BS) and users, or deployed on the BS. Compared with traditional RIS application scenarios, US-RIS is a hardware technology that beamforming the signals that penetrate them and is used to build large-scale arrays on the user side.

In this paper, we use alternating update and SCA methods to solve the optimal energy efficiency of the US-RIS-aided uplink communication system. Thereby the optimal layer of US-RIS is found. Simulation results demonstrate that the UE's internal circuit consumption and the number of elements can affect the energy efficiency, and the optimal layer changes accordingly.

2 System Model

In Fig. 1, a US-RIS-aided uplink communication system is considered. The system has one UE equipped with M transmit antennas and one BS with N receive antennas. The US-RIS consists of L layers and each layer has K_l passive elements, $l = [1, 2, \dots, L]$.

The phase shift (PS) matrix of US-RIS's l -th layer is defined as

$$\Theta_l = \text{diag}(\boldsymbol{\theta}_l) = \text{diag}[(\theta_{l,1}, \dots, \theta_{l,K_l})^T], \quad (1)$$

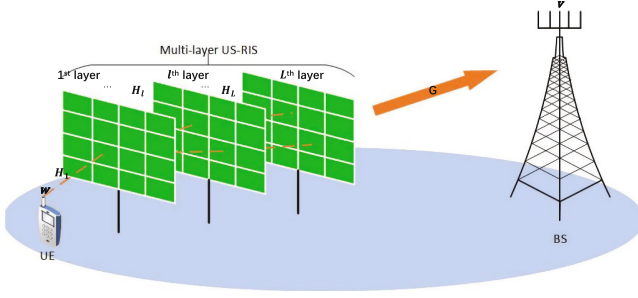


Fig. 1. System model

where, $\theta_{l,k} = e^{j\varphi}$ is k -th element's phase shift of the l -th layer and $\varphi \in [-\pi, \pi]$. Let $s \in \mathbb{C}$ denote the transmit signal for uplink transmission with $\mathbb{E}\{|s|^2\} = 1$. The received vector $\mathbf{y} \in \mathbb{C}^{N \times 1}$ at the BS can be written as

$$\mathbf{y} = \mathbf{G}^H \left(\prod_{l=1}^L \kappa \Theta_l \mathbf{H}_l \right) \mathbf{w} s + \mathbf{n}, \quad (2)$$

where $\mathbf{G} \in \mathbb{C}^{K_L \times N}$ denotes the channel between the l -th layer of US-RIS and BS. $\mathbf{H}_1 \in \mathbb{C}^{K_1 \times M}$ denotes the channel between US-RIS's 1-st layer and UE. $\mathbf{H}_l \in \mathbb{C}^{K_l \times K_{l-1}}, l = [2, \dots, L]$ denotes the channel between the $(l-1)$ -th layer and the l -th layer of US-RIS. In addition, κ is the loss coefficient when the signal transmit through every layer. $\mathbf{w} \in \mathbb{C}^{M \times 1}$ is the uplink beamforming (ULB) vector. \mathbf{n} is a Gaussian White noise vector with zero mean and $\sigma^2 \mathbf{I}_N$ variance matrix. Therefore, the transmission signal which is combined by the BS can be expressed as

$$z = \mathbf{v}^H \mathbf{G}^H \left(\prod_{l=1}^L \kappa \Theta_l \mathbf{H}_l \right) \mathbf{w} s + \mathbf{v}^H \mathbf{n}, \quad (3)$$

where $z = \mathbf{v}^H \mathbf{y}$. The normalized receiving combining (RC) vector $\mathbf{v} \in \mathbb{C}^{N \times 1}$ is used here. Then, SNR is expressed as the following form:

$$SNR = \frac{|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa \Theta_l \mathbf{H}_l) \mathbf{w}|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2}. \quad (4)$$

Thus, the achieved data rate is

$$R = \log_2 \left(1 + \frac{|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa \Theta_l \mathbf{H}_l) \mathbf{w}|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2} \right). \quad (5)$$

Note that the rate in (5) is affected by the RC vector \mathbf{v} , US-RIS PS matrices $\Theta_1, \dots, \Theta_L$, ULB vector \mathbf{w} . Hence, the problem of energy efficient can be formulated as

$$\max_{\mathbf{v}, \Theta_1, \dots, \Theta_L, \mathbf{w}} \frac{R}{\|\mathbf{w}\|_2^2 + P_i}, \quad (6a)$$

$$\text{s.t. } 0 \leq \|\mathbf{w}\|_2^2 \leq P_{max}, \quad (6b)$$

$$R \geq r_{min}, \quad (6c)$$

$$|\theta_{l,k}| = 1, \quad (6d)$$

where P_i is the internal power circuit consumption of UE. (6b) is the power constraint of UE. (6c) is the constraint to ensure the minimum rate at the receiving end, namely, quality of service (QoS). (6d) is the range of phase shift of the k -th element of the l -th layer.

Algorithm 1. Alternating Update with SCA

Input: Channel matrices \mathbf{G} and $\mathbf{H}_1, \dots, \mathbf{H}_L$; loss factor κ ; $t = 1$; maximum transmit power P_{max} and internal power circuit consumption P_i .

Output: Maximized energy efficient τ .

- 1: **while** τ does not converge **do**
 - 2: Update \mathbf{v}^{opt} by (11);
 - 3: Update $\Theta_1^{opt}, \dots, \Theta_L^{opt}$ by (15);
 - 4: Update $\mathbf{w}^{(t)}$ and $\tau^{(t)}$ by solve problem (28);
 - 5: Set $t = t + 1$;
 - 6: **end while**
 - 7: **return** $\mathbf{v}^{opt}, \Theta_1^{opt}, \dots, \Theta_L^{opt}$ and $\mathbf{w}^{(t)}$.
-

3 The Optimal Energy Efficient Design

To address the formulate non-convex problem, we design an algorithm which can efficiently solve the problem. We want to find the maximum energy efficient of US-RIS-aided uplink communication system according to the beamforming. The proposed algorithm is used to settle the joint optimization problem (6). Given the input channel matrices $\mathbf{H}_1, \dots, \mathbf{H}_L$, \mathbf{G} and loss factor κ . By alternately solve the subproblem, the \mathbf{v}^{opt} is the first update with Rayleigh quotient problem, and then $\Theta_1^{opt}, \dots, \Theta_L^{opt}$ is update. Further we use successive convex approximation(SCA) method to solve \mathbf{w} .

3.1 Receiving Combining Vector Design

With other variables fixed, problem (6) can be transformed into the following equivalent form:

$$\begin{aligned} & \max_{\mathbf{v}, \|\mathbf{v}\|_2^2=1} \frac{|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w}|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2} \\ & = \max_{\mathbf{v}, \|\mathbf{v}\|_2^2=1} \frac{\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w} \mathbf{w}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l)^H \mathbf{G} \mathbf{v}}{\|\mathbf{v}^H\|_2^2 \sigma^2}. \end{aligned} \quad (7)$$

Let $\mathbf{U} = \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w} \mathbf{w}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l)^H \mathbf{G}$. Obviously, \mathbf{U} is a Hermitian matrix, and equation (7) can be written as

$$\max_{\mathbf{v}} \frac{\mathbf{v}^H \mathbf{U} \mathbf{v}}{\|\mathbf{v}^H\|_2^2 \sigma^2}, \quad (8a)$$

$$\text{s.t. } \|\mathbf{v}\|_2^2 = 1, \quad (8b)$$

At this point, the problem is transformed into the Rayleigh quotient problem. And (8) can be formulated to

$$R(\mathbf{U}, \mathbf{v}) = \frac{\mathbf{v}^H \mathbf{U} \mathbf{v}}{\mathbf{v}^H \mathbf{v}}. \quad (9)$$

The generalized Rayleigh quotient function (9) is satisfied the relationship as follows

$$\lambda_{min} \leq \frac{\mathbf{v}^H \mathbf{U} \mathbf{v}}{\mathbf{v}^H \mathbf{v}} \leq \lambda_{max}, \quad (10)$$

where λ_{max} is the maximum eigenvalue of matrix \mathbf{U} and λ_{min} is the minimum eigenvalue of matrix \mathbf{U} . So, the \mathbf{v}^{opt} can be expressed as

$$\mathbf{v}^{opt} = \lambda_{max}(\mathbf{U}). \quad (11)$$

3.2 Phase Shift Matrix Design

With other variables fixed, and only considering Θ_l , problem (6) can be equivalently written as

$$\max_{\Theta_1, \dots, \Theta_L} R, \quad (12a)$$

$$\text{s.t. } R \geq r_{min}, \quad (12b)$$

$$|\theta_{l,k}| = 1. \quad (12c)$$

We can define

$$\psi_{(i,j)} = \prod_{l=i}^j \kappa_l \Theta_l \mathbf{H}_l, \quad (i \in L, \quad j \in L). \quad (13)$$

The SNR can be equivalently written as

$$\begin{aligned} SNR &= \frac{|\mathbf{v}^H \mathbf{G}^H \boldsymbol{\psi}_{(l+1,L)} \boldsymbol{\psi}_{(1,l)} \mathbf{w}|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2} \\ &= \frac{|\mathbf{v}^H \mathbf{G}^H \boldsymbol{\psi}_{(l+1,L)} \text{diag}(\mathbf{H}_l \boldsymbol{\psi}_{(1,l-1)} \mathbf{w}) \boldsymbol{\theta}_l|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2}. \end{aligned} \quad (14)$$

Then, with the constraint (12b) and (12c), the optimal phase can be written by

$$\boldsymbol{\theta}_l^{opt} = \begin{cases} \exp\left(j \arg\left(\text{diag}\left(\mathbf{H}_l \boldsymbol{\psi}_{(1,l-1)} \mathbf{w}\right)^H \boldsymbol{\psi}_{(l+1,L)}^H \mathbf{G} \mathbf{v}\right)\right), R \geq r_{min} \\ \text{No solution, } R < r_{min} \end{cases} \quad (15)$$

3.3 Uplink Beamforming Vector Design

To make the problem (6) convex, we use a slack variable τ to transform (6) equivalently as follows

$$\max_{\mathbf{w}, \tau} \tau, \quad (16a)$$

$$\text{s.t. } \frac{R}{\|\mathbf{w}\|_2^2 + P_i} \geq \tau, \quad (16b)$$

$$0 \leq \|\mathbf{w}\|_2^2 \leq P_{max}, \quad (16c)$$

$$R \geq r_{min}. \quad (16d)$$

Since (16b) is non-convex, we introduce a slack variable α to make the constraint (16b) convex. Therefore, the power constraint can be expressed as

$$\|\mathbf{w}\|_2^2 + P_i \leq \alpha, \quad (17)$$

and constraint (16b) transform into the following form

$$R \geq \tau \alpha. \quad (18)$$

But the power constraint (18) is still non-convex, use a new variable β to relax (18), which is written as

$$\log_2(\beta) \geq \tau \alpha, \quad (19a)$$

$$1 + SNR \geq \beta. \quad (19b)$$

Then, we use another variable γ to make constraint (19a) convex, which can be expressed as

$$\gamma \geq \tau \alpha, \quad (20a)$$

$$\beta \geq 2^\gamma. \quad (20b)$$

The relationship between these series of slack variables can be described as

$$\log_2(1 + SNR) \geq \log_2(\beta) \geq \gamma \geq \tau\alpha, \quad (21)$$

Due to the fractional form of SNR, a variable Ω can be introduced. Thus, (19b) can be equivalently written as

$$|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w}|^2 \geq (\beta - 1)\Omega, \quad (22a)$$

$$\|\mathbf{v}^H\|_2^2 \sigma^2 \leq \Omega. \quad (22b)$$

In addition, we use a slack variable δ , (19a) can be expressed as

$$\delta^2 \geq (\beta - 1)\Omega, \quad (23a)$$

$$|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w}|^2 \geq \delta^2. \quad (23b)$$

(19a) and (23a) are non-convex, we use the first-order Taylor to approximate non-convex constraints. These constraints can be expressed as

$$\gamma \geq \tau^{(t)} \alpha^{(t)} + \alpha^{(t)} (\tau - \tau^{(t)}) + \tau^{(t)} (\alpha - \alpha^{(t)}), \quad (24)$$

$$2\delta^{(t)}\delta - (\delta^{(t)})^2 \geq (\beta^{(t)} - 1)\Omega^{(t)} + \Omega^{(t)} (\beta - \beta^{(t)}) + (\beta^{(t)} - 1)(\Omega - \Omega^{(t)}), \quad (25)$$

where t represents the t -th iteration. Then the non-convex constraint (16d) can be shown as

$$\frac{|\mathbf{v}^H \mathbf{G}^H (\prod_{l=1}^L \kappa_l \Theta_l \mathbf{H}_l) \mathbf{w}|^2}{\|\mathbf{v}^H\|_2^2 \sigma^2} \geq 2^{r_{min}} - 1, \quad (26)$$

where r_{min} is a constant. Therefore, this constraint can be restated as

$$2\delta^{(t)}\delta - (\delta^{(t)})^2 \geq (2^{r_{min}} - 1)(\|\mathbf{v}^H\|_2^2 \sigma^2). \quad (27)$$

Constraints (27) is convex. The convex optimization problem can be described in following form:

$$\max_{\mathbf{w}, \tau, \alpha, \beta, \gamma, \Omega, \delta} \tau, \quad (28a)$$

$$\text{s.t. } (16c), (17), (20b), (24), (25), (27). \quad (28b)$$

Hence, we can solve the problem using the CVX.

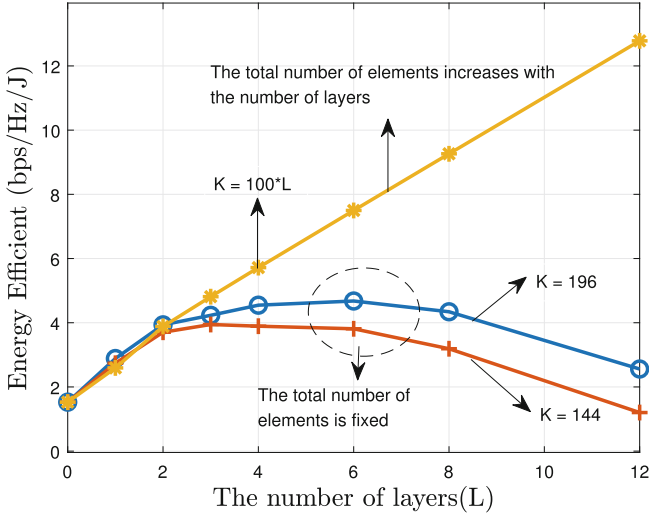


Fig. 2. Energy efficient versus US-RISs' layers

4 Simulation and Discussion

In this section, we conduct some experiments to quantify the energy efficient of the US-RIS-aided uplink communication system after increasing or decreasing the number of layers of US-RIS with and without changing the total number of elements. We then discuss the effect of different internal circuits consumption P_i on energy efficiency. Let λ represent the wavelength of the uplink signal. The side length of each element is $\frac{\lambda}{2}$, and each element is tightly pressed together without spacing. We assume that user having 2 transmitting antennas, and the BS having receiving 8 antennas. The distance between each antenna is $\frac{\lambda}{2}$. We assume that the distance between UE and the first layer of US-RIS is 0.02 m, and the distance between UE and BS is 20 m. Among them, the distance between each layer of US-RIS's layers is 0.02 m. The frequency of uplink transmission signal is configured to 2.5 GHz. The loss coefficient is set to 0.8. The variance of noise is set to $\sigma^2 = -60$ dB. Furthermore, we assume that the uplink communication system knows perfect channel state information (CSI). We initialize the phases of all RIS to random phase values in the feasible set. Thus, we can find the optimal layer according to different conditions.

As shown in Fig. 2, energy efficiency is related to the number of transmissive elements. The total numbers of elements are set as $K = 196$ and $K = 144$. P_i is set to -30 dB and P_{max} is set to 3 dB. The case that the total number of elements increases with the number of layers are considered. In Fig. 2, we can observe that the energy efficiency of the number of elements increasing with the number of layers is significantly better than that of the number of fixed elements for $L \geq 3$. For $L = 12$, the gap from the fixed element number of $K = 192$ to the total number without limited is 10.23 bps/Hz/J. In addition, we can find

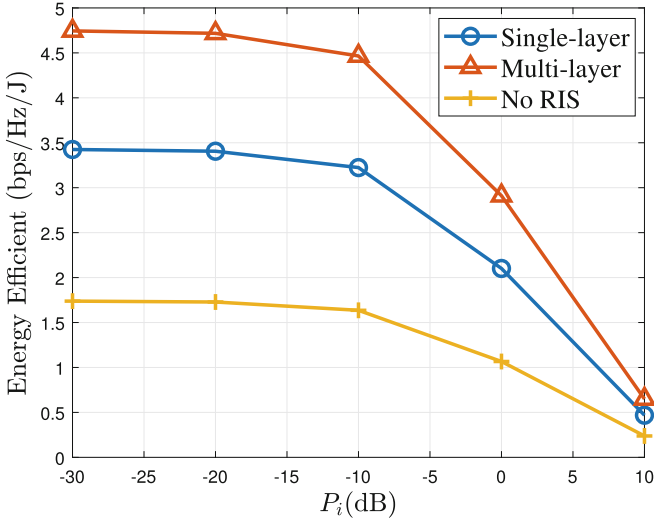


Fig. 3. Energy efficient versus P_i

the optimal number of layers. For $K = 196$, the optimal layer is $L = 6$, and its energy efficient is 4.68 bps/Hz/J. For $K = 144$ the optimal layer is $L = 2$, and its energy efficient is 3.95 bps/Hz/J. An increase in the number of elements can be found to increase the energy efficiency of the system. With a fixed total number of elements, it can also be observed that the number of elements affects the optimal layer for US-RIS.

Figure 3 reveals that energy efficiency is affected by internal circuit consumption. The multi-layer and single-layer cases of US-RIS of the uplink communication system are compared, and the curve of No RIS is used as a baseline. The number of element is set to 192. P_i is set from -30 dB to 10 dB. P_{max} is set to 2 dB. The three curves for single layer, multi-layer and no RIS can see that the energy efficiency decreases with the increase of the P_i . For P_i from -30 dB to 0 dB, the energy efficiency drops more slowly. For P_i from -10 dB to 10 dB, the energy efficiency drops drastically. The efficiency is close to 0 for $P_i = 10$ dB. Energy efficiency decreases as P_i increases. After P_i is more than P_{max} , the energy efficiency decreases slowly and approaches zero.

Figure 4 depicts the convergence of the proposed algorithm under existing simulation conditions. Case of No-RIS as a comparison, the results clearly show that the proposed algorithm for multi-layer US-RIS converges at the fourth iteration. And the proposed algorithm for sigle-layer US-RIS converges at the first iteration.

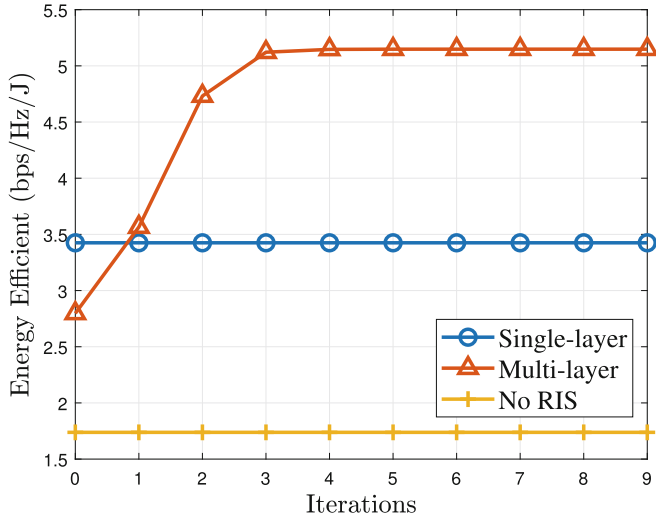


Fig. 4. Energy efficient against iterations

5 Conclusion

In this paper, we formulated a non-convex problem. To address this problem, we subsequently devised an algorithm using alternating update and SCA methods. This problem is divided into three sub-problems and solved separately. To solve the energy efficiency of the system, we alternatingly update the solutions of the three subproblems. Finally, the optimal layer of US-RIS is found by energy efficiency. The simulation results show that the optimal layer of US-RIS has been found. The optimal layer is affected by the addition or subtraction of elements, and increasing the number of layers does not necessarily lead to an increase in energy efficiency.

References

1. Yu, L., et al.: Sparse code multiple access for 6G wireless communication networks: recent advances and future directions. *IEEE Commun. Stan. Mag.* **5**(2), 92–99 (2021)
2. Cai, D., Ding, Z., Fan, P., Yang, Z.: On the performance of NOMA with hybrid ARQ. *IEEE Trans. Veh. Technol.* **67**(10), 10033–10038 (2018)
3. Pham, Q.V., et al.: A survey of multi-access edge computing in 5G and beyond: fundamentals, technology integration, and state-of-the-art. *IEEE Access* **8**, 116974–117017 (2020)
4. Tariq, F., Khandaker, M.R.A., Wong, K.-K., Imran, M.A., Bennis, M., Debbah, M.: A speculative study on 6G. *IEEE Wireless Commun.* **27**(4), 118–125 (2020)
5. Cai, D., Xu, Y., Fang, F., Ding, Z., Fan, P.: On the impact of time-correlated fading for downlink NOMA. *IEEE Trans. Commun.* **67**(6), 4491–4504 (2019)

6. Dang, S., Amin, O., Shihada, B., Alouini, M.S.: What should 6G be? *Nat. Electron.* **3**(1), 20–29 (2020)
7. Jiao, S., Fang, F., Zhou, X., Zhang, H.: Joint beamforming and phase shift design in downlink UAV networks with IRS-assisted NOMA. *J. Commun. Inf. Netw.* **5**(2), 138–149 (2020)
8. ElMossallamy, M.A., Zhang, H., Song, L., Seddik, K.G., Han, Z., Li, G.Y.: Reconfigurable intelligent surfaces for wireless communications: principles, challenges, and opportunities. *IEEE Trans. Cognit. Commun. Netw.* **6**(3), 990–1002 (2020)
9. Wu, Q., Zhang, R.: Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network. *IEEE Commun. Mag.* **58**(1), 106–112 (2019)
10. Pan, C., et al.: Reconfigurable intelligent surfaces for 6G systems: principles, applications, and research directions. *IEEE Commun. Mag.* **59**(6), 14–20 (2021)
11. Dai, L., et al.: Reconfigurable intelligent surface-based wireless communications: antenna design, prototyping, and experimental results. *IEEE Access* **8**, 45913–45923 (2020)
12. Zhang, Z., Dai, L.: A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network. *IEEE Trans. Signal Process.* **69**, 4085–4101 (2021)
13. Huang, C., Zappone, A., Alexandropoulos, G.C., Debbah, M., Yuen, C.: Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Trans. Wireless Commun.* **18**(8), 4157–4170 (2019)
14. Zeng, S., et al.: Reconfigurable intelligent surfaces in 6G: reflective, transmissive, or both? *IEEE Commun. Lett.* **25**(6), 2063–2067 (2021)
15. Basar, E., Di Renzo, M., De Rosny, J., Debbah, M., Alouini, M.S., Zhang, R.: Wireless communications through reconfigurable intelligent surfaces. *IEEE Access* **7**, 116753–116773 (2019)
16. Wang, T., Fang, F., Ding, Z.: An SCA and relaxation based energy efficiency optimization for multi-user RIS-assisted NOMA networks. *IEEE Trans. Veh. Technol.* (2022)
17. Li, Y., et al.: Transmission-type 2-bit programmable metasurface for single-sensor and single-frequency microwave imaging. *Sci. Rep.* **6**(1), 1–8 (2016)
18. Liu, K., Zhang, Z., Dai, L., Hanzo, L.: Compact user-specific reconfigurable intelligent surfaces for uplink transmission. *IEEE Trans. Commun.* **70**(1), 680–692 (2021)



A Sophisticated Anti-jamming Strategy for a Joint Radar and Communication System

Andrey Garnaev^(✉) and Wade Trappe

WINLAB, Rutgers University, North Brunswick, USA
garnaev@yahoo.com, trappe@winlab.rutgers.edu

Abstract. In this paper, we consider the problem of determining how a joint (dual) radar and communication system should divide its effort between supporting its radar and communication tasks in the presence of a jammer that wants to obstruct the system's work by means of jamming. The system, besides the basic objective consisting of two tasks (a) to communicate with a receiver and (b) to track a radar target through the reflections witnessed at the system, also has the secondary objective to achieve the basic objective in a manner that is as unpredictable as possible to the jammer. The signal to interference and noise ratio (SINR) of the radar and communication's SINR are considered as the metrics that reflect the radar and communication tasks, respectively. The entropy associated with a system's strategy to switch between two tasks is considered as a metric that reflects unpredictability of its strategy for the jammer. We model this problem by a Bayesian game for a scenario where the system is at a disadvantage to access information about environmental parameters relative to the jammer. The established uniqueness of the equilibrium reflects stability of the designed anti-jamming strategy, even in such a disadvantageous situation for the system.

Keywords: Nash equilibrium · Radar · Communication · Jamming

1 Introduction

There has been recent interest in enabling radar and communication systems to co-exist in the same frequency bands in order to allow spectrum to be utilized more efficiently [17]. This has given rise to a significant amount of research on methods for spectrum sharing between the two systems. One approach to achieve this is to formulate waveform design using Orthogonal Frequency Division Multiplexing (OFDM) signals and then optimally allocate the subcarriers [12, 18]. Radar waveform design for controlled interference is considered in [3, 4], while the cooperative design of the two systems was explored in [5, 14]. In this paper, we consider a dual purpose communication-radar system that employs OFDM style waveforms and explore the *complementary aspect* where the system besides the basic objective consisting of two tasks (a) to communicate with a receiver

and (b) to track a radar target through the reflections witnessed at the system, also has the secondary objective to achieve the basic objective in a manner that is as unpredictable as possible to the jammer. We model this problem by a non-zero sum (Bayesian) game. The radar's SINR and communication's SINR are considered as the metrics that reflect corresponding tasks of the basic objective. The entropy of the system's strategy for switching between its two tasks of basic objective is considered as the metric that reflects its unpredictability, i.e., its secondary objective. The problem is modeled by a non-zero sum game and the equilibrium is found in closed form. Its uniqueness is proven, which reflects stability of the designed anti-jamming strategy even in the case where the system is at a serious disadvantage when compared with the jammer in terms of the information possessed regarding the jamming fading gains. Specifically, we consider that the system has only statistical information, meanwhile the jammer has complete information about the channel gains.

2 Basic Model of a Dual Purpose Communication-Radar System

Let us consider an operational scenario involving a dual purpose communication-radar system that transmits communication signals only, and uses the received reflections of those signals off targets for target tracking in presence of an jammer (Fig. 1). The system wants to support two different tasks: (a) to communication with a receiver, and (b) to track a radar target through the reflections witnessed at the system. In order to support these two tasks, the system uses a spectrum band that is modeled as consisting of n adjacent sub-channels, which may be associated with n different subcarriers. In this paper we employ a transmission scheme like OFDM, as considered in [11]. With each of these n different subcarriers, two different (fading) channel gains are associated. Specifically, we let $h_{\mathbb{R},i}$, $i \in \mathcal{N} \triangleq \{1, \dots, n\}$, correspond to the i -th radar channel gain associated with the round-trip effect of the transmitted signal, reflected off the radar target, and received at the system, while $h_{\mathbb{C},i}$ denotes the i -th channel gain associated with the i -th communication subcarrier at the receiver.

Let $\mathbf{P} = (P_1, \dots, P_n)$ be power-allocation strategy for the system, where P_i , $i \in \mathcal{N}$, is the power assigned for transmitting on subcarrier i , and $\sum_{i \in \mathcal{N}} P_i = \bar{P}$, where \bar{P} is the total power budget.

We assume that to avoid interference at each moment the system focuses only on one of two tasks. We call such focusing on a task by implementing the corresponding *mode*. Thus, at each moment the system can implement one of two modes:

- (i) *communication mode*, denoted by \mathbb{C} , to focus on the communication's task,
- (ii) *radar mode*, denoted by \mathbb{R} , to focus on the radar's task.

In order to unify the examination of radar and communication metrics, we note that radar detection/tracking and communication throughput are both closely related to the associated signal-to-interference-plus-noise ratio (SINR)

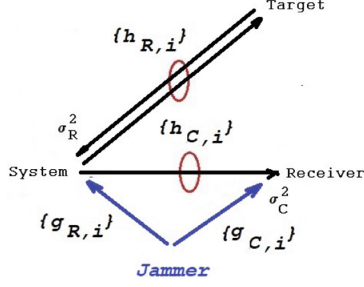


Fig. 1. Dual purpose communication-radar system, target, receiver and jammer

as witnessed at the appropriate recipient. Let the radar SINR and communication SINR be given by $\sum_{i=1}^n h_{\mathbb{R},i}P_i/\sigma_{\mathbb{R}}^2$ and $\sum_{i=1}^n h_{\mathbb{C},i}P_i/\sigma_{\mathbb{C}}^2$, respectively, with $\sigma_{\mathbb{R}}^2$ and $\sigma_{\mathbb{C}}^2$ are corresponding background noise powers. For the communication objective, the SINR is used as the payoff function for two reasons: first, it is easily linearized; and, second, for a low SINR regime, SINR is an approximation to throughput. For the radar objective, SINR is used as the payoff function since it is closely related to the associated detection metrics [16].

3 Jammer

Let a *jammer* be present, and it is located close to the receiver (Fig. 1). The jammer seeks to introduce hostile interference to disrupt the functionality of the system. Its strategy is a power allocation vector $\mathbf{J} = (J_1, \dots, J_n)$, where J_i , $i \in \mathcal{N}$, is the power assigned for jamming subcarrier i , and $\sum_{i \in \mathcal{N}} J_i = \bar{J}$ where \bar{J} is the total power budget allocated for jamming. We assume that the jammer's signal will not reach the target. Under a jamming attack, the communication's SINR and radar's SINRs are given by:

$$\text{SINR}_{\mathbb{R}}(\mathbf{P}, \mathbf{J}) = \sum_{i \in \mathcal{N}} \frac{h_{\mathbb{R},i}P_i}{\sigma_{\mathbb{R}}^2 + g_{\mathbb{R},i}J_i} \quad (1)$$

and

$$\text{SINR}_{\mathbb{C}}(\mathbf{P}, \mathbf{J}) = \sum_{i \in \mathcal{N}} \frac{h_{\mathbb{C},i}P_i}{\sigma_{\mathbb{C}}^2 + g_{\mathbb{C},i}J_i}, \quad (2)$$

where $g_{\mathbb{R},i}$ are fading channels gains between the jammer and the radar, and $g_{\mathbb{C},i}$ are fading gains between the jammer and the communication receiver.

Let the system work in a fixed mode m , $m \in \{\mathbb{C}, \mathbb{R}\}$ and this mode is known to the jammer. Let the system and jammer implement strategies \mathbf{P} and \mathbf{J} , respectively. Then, $\text{SINR}_m(\mathbf{P}, \mathbf{J})$ is the payoff to system in mode m , and it can be considered as the cost function for the jammer. Thus, for a fixed mode m ,

this is a zero-sum game with $\text{SINR}_m(\mathbf{P}, \mathbf{J})$ as system's payoff. Let us denote this zero-sum game by \mathfrak{G}_m .

For each fixed mode $m \in \{\mathbb{C}, \mathbb{R}\}$, we look for a (Nash) equilibrium in this game \mathfrak{G}_m . Thus, for such a pair of strategies $(\mathbf{P}_m, \mathbf{J}_m)$ that the following inequalities hold:

$$\text{SINR}_m(\tilde{\mathbf{P}}, \mathbf{J}_m) \leq \text{SINR}_m(\mathbf{P}_m, \mathbf{J}_m) \leq \text{SINR}_m(\mathbf{P}_m, \tilde{\mathbf{J}}) \text{ for all } (\tilde{\mathbf{P}}, \tilde{\mathbf{J}}). \quad (3)$$

The equilibrium strategies $(\mathbf{P}_m, \mathbf{J}_m)$ in game \mathfrak{G}_m can be found via [1].

4 Mode Selection

Now suppose the system can choose between radar and communication mode, and, then, implement the optimal power allocation strategy for these modes $\mathbf{P}_{\mathbb{R}}$ and $\mathbf{P}_{\mathbb{C}}$, respectively. The jammer does not know what mode the system implements. Thus, a dilemma for the jammer arises: the best response to each of the system's power allocation strategies $\mathbf{P}_{\mathbb{R}}$ or $\mathbf{P}_{\mathbb{C}}$ to implement. Specifically, to implement power allocation strategy $\mathbf{J}_{\mathbb{R}}$ or $\mathbf{J}_{\mathbb{C}}$, respectively. Meanwhile, in its turn, the system does not know which decision will be made by the jammer. Thus, a dilemma also arises for the system: the best response to each of jammer's power allocation strategies $\mathbf{P}_{\mathbb{R}}$ or $\mathbf{P}_{\mathbb{C}}$ to implement, namely, to apply power allocation strategy $\mathbf{J}_{\mathbb{R}}$ or $\mathbf{J}_{\mathbb{C}}$, respectively. This scenario leads to the following payoff matrix \mathcal{M} , where the rows are the system's strategies and the columns are the jammer's strategies:

$$\mathcal{M} = \begin{array}{c} \begin{array}{cc} & \mathbb{R} & \mathbb{C} \\ \mathbb{R} & A & a \\ \mathbb{C} & b & B \end{array} \end{array} \quad (4)$$

with

$$A \triangleq \text{SINR}_{\mathbb{R}}(\mathbf{P}_{\mathbb{R}}, \mathbf{J}_{\mathbb{R}}), \quad a \triangleq \text{SINR}_{\mathbb{R}}(\mathbf{P}_{\mathbb{R}}, \mathbf{J}_{\mathbb{C}}), \quad (5)$$

$$B \triangleq \text{SINR}_{\mathbb{C}}(\mathbf{P}_{\mathbb{C}}, \mathbf{J}_{\mathbb{C}}) \text{ and } b \triangleq \text{SINR}_{\mathbb{C}}(\mathbf{P}_{\mathbb{C}}, \mathbf{J}_{\mathbb{R}}). \quad (6)$$

Note that, by (3), without loss of generality we could assume that

$$A < a \text{ and } B < b. \quad (7)$$

Let $\mathbf{x} = (x, 1 - x)$ be randomized (mixed) strategies [6] for the system, where x and $1 - x$ are probabilities that the system is in mode \mathbb{R} and \mathbb{C} , respectively. Meanwhile, $\mathbf{y} = (y, 1 - y)$, be randomized (mixed) strategies for the jammer, where y and $1 - y$ are probabilities that the jammer is in mode \mathbb{R} and \mathbb{C} , respectively. Then, if the system and the jammer implement strategies \mathbf{x} and \mathbf{y} , respectively, the expected system's SINR is given as follows

$$\begin{aligned} \text{SINR}_S(\mathbf{x}, \mathbf{y}) &= \mathbf{x} \mathcal{M} \mathbf{y}^T \\ &= Axy + ax(1 - y) + b(1 - x)y + B(1 - x)(1 - y). \end{aligned} \quad (8)$$

4.1 Sophisticated System

The (sophisticated) system wants to find a trade-off between two objectives:

- (i) *the basic objective*, to maximize the expected system's SINR,
- (ii) *the secondary objective*, to achieve this maximum in the most unpredictable way for the jammer.

As a metric for the system to confuse the jammer, we consider the entropy of its strategy, i.e.,

$$H(\mathbf{x}) = -x \ln(x) - (1-x) \ln(1-x). \quad (9)$$

The payoff for such (sophisticated) system is taken as a weighted sum of the expected system's SINR and the entropy of its strategy, i.e.,

$$v_S(\mathbf{x}, \mathbf{y}) = w_S \text{SINR}_S(\mathbf{x}, \mathbf{y}) + w_E H(\mathbf{x}), \quad (10)$$

where w_S and w_E are non-negative weighting coefficients.

Note that for any scale positive parameter C , the payoff $C v_S(\mathbf{x}, \mathbf{y})$ achieves its maximum at the same strategy \mathbf{x} . That is why, without loss of generality we can assume that w_S and w_E are normalized non-negative weighting coefficients. Specifically,

$$w_S \triangleq 1 - w \text{ and } w_E \triangleq w, \text{ where } w \in [0, 1]. \quad (11)$$

This and (10) imply that

$$v_S(\mathbf{x}, \mathbf{y}) = (1-w) \text{SINR}_S(\mathbf{x}, \mathbf{y}) + w H(\mathbf{x}). \quad (12)$$

Note that, in [8] we used entropy metric to design a sophisticated adversary's strategy for illegal sneaking into protected bandwidth, and, in [7,9], to design sophisticated anti-eavesdropping strategies,

Meanwhile, jammer wants to minimize the expected system's SINR. Thus, payoff to the jammer is negative of the expected system's SINR, i.e.,

$$v_J(\mathbf{x}, \mathbf{y}) = -\text{SINR}_S(\mathbf{x}, \mathbf{y}). \quad (13)$$

5 Incomplete Information

In this section, we consider a more practical case, in which there is incomplete information about some of the parameters involved. In particular, we assume that the jammer knows all of the parameters perfectly, as this is the most dangerous case for the system. Meanwhile, the system knows all of the parameters except the channel gains for the jammer, for which it has limited statistical information. Specifically, the system only knows that the jammer's channel gains can have values $(g_{\mathbb{R},i,k}, g_{\mathbb{C},i,k})$, $i \in \mathcal{N}$ with probability α_k , where $k \in \mathcal{K} \triangleq \{1, \dots, K\}$. In other words, the jammer's channel gains can be in one of K states, and the

system does not know which state occurs (Fig. 2). Note that if the jammer’s channel gains are in state k , then the radar’s SINR and communication’s SINR are give as follows:

$$\text{SINR}_{\mathbb{R},k}(\mathbf{P}, \mathbf{J}) = \sum_{i \in \mathcal{N}} \frac{h_{\mathbb{R},i} P_i}{\sigma_{\mathbb{R}}^2 + g_{\mathbb{R},i,k} J_i} \tag{14}$$

and

$$\text{SINR}_{\mathbb{C},k}(\mathbf{P}, \mathbf{J}) = \sum_{i \in \mathcal{N}} \frac{h_{\mathbb{C},i} P_i}{\sigma_{\mathbb{C}}^2 + g_{\mathbb{C},i,k} J_i}. \tag{15}$$

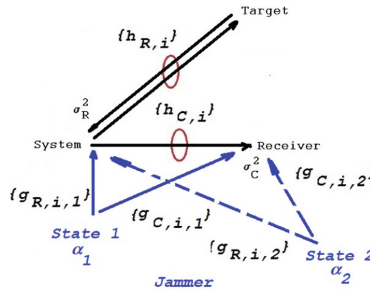


Fig. 2. Dual purpose communication- radar system, target, receiver and jammer, which channel gains can be in one of two states, i.e., $\mathcal{K} = \{1, 2\}$.

Similar to (3), for each fixed mode $m \in \{\mathbb{C}, \mathbb{R}\}$ and jammer’s channel state $k \in \mathcal{K}$ let $\mathbf{P}_{m,k}$ and $\mathbf{J}_{m,k}$ be equilibrium strategies of the system and jammer, respectively, in zero sum game with $\text{SINR}_{m,k}$ as system’s payoff. Further, if channel state k occurs, and this state is known to both players, the system and jammer can choose mode $m \in \{\mathbb{R}, \mathbb{C}\}$ and $\tilde{m} \in \{\mathbb{R}, \mathbb{C}\}$, respectively, and implement the equilibrium strategies $\mathbf{P}_{m,k}$ and $\mathbf{J}_{\tilde{m},k}$, associated with their choices independently from each other. This leads to the following payoff matrix \mathcal{M}_k to the system:

$$\mathcal{M}_k = \begin{matrix} & \mathbb{R} & \mathbb{C} \\ \begin{matrix} \mathbb{R} \\ \mathbb{C} \end{matrix} & \begin{pmatrix} A_k & a_k \\ b_k & B_k \end{pmatrix} \end{matrix} \tag{16}$$

with

$$A_k \triangleq \text{SINR}_{\mathbb{R},k}(\mathbf{P}_{\mathbb{R},k}, \mathbf{J}_{\mathbb{R},k}), \quad a_k \triangleq \text{SINR}_{\mathbb{R},k}(\mathbf{P}_{\mathbb{R},k}, \mathbf{J}_{\mathbb{C},k}), \tag{17}$$

$$B_k \triangleq \text{SINR}_{\mathbb{C},k}(\mathbf{P}_{\mathbb{C},k}, \mathbf{J}_{\mathbb{C},k}) \text{ and } b_k \triangleq \text{SINR}_{\mathbb{C},k}(\mathbf{P}_{\mathbb{C},k}, \mathbf{J}_{\mathbb{R},k}). \tag{18}$$

Then, following (3), we have that

$$A_k < a_k \text{ and } B_k < b_k. \tag{19}$$

Let us denote as a *k-type jammer*, the jammer if its fading gains occur in state k . Also, $\mathbf{y}_k = (y_k, 1 - y_k)$, be mixed strategies for the k -type jammer, where y_k and $1 - y_k$ are probabilities that k -type jammer is in mode \mathbb{R} and \mathbb{C} , respectively. Since the system does not know which fading gain state of the jammer occurs, the system strategy is $\mathbf{x} = (x, 1 - x)$, where x and $1 - x$ are probabilities that the system is in mode \mathbb{R} and \mathbb{C} , respectively.

Then, the expected system's SINR is given by (20) below if the system and jammer, dependently on its type, implement strategies \mathbf{x} and $\mathbf{y}_1, \dots, \mathbf{y}_K$, respectively:

$$\text{SINR}_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K) = \sum_{k \in \mathcal{K}} \alpha_k \mathbf{x} \mathcal{M}_k \mathbf{y}_k^T. \quad (20)$$

By (12) and (20), payoffs to sophisticated system k -type jammer are given, respectively as follows:

$$v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K) = (1 - w) \sum_{k \in \mathcal{K}} \alpha_k \mathbf{x} \mathcal{M}_k \mathbf{y}_k^T + wH(\mathbf{x}), \quad (21)$$

$$v_{J,k}(\mathbf{x}, \mathbf{y}_k) = -\mathbf{x} \mathcal{M}_k \mathbf{y}_k^T, \quad k \in \mathcal{K}. \quad (22)$$

We look for a Nash equilibrium. i.e., for such feasible strategies $(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)$ that any feasible strategies $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_K)$, the following inequalities hold:

$$v_S(\tilde{\mathbf{x}}, \mathbf{y}_1, \dots, \mathbf{y}_K) \leq v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K), \quad (23)$$

$$v_{J,k}(\mathbf{x}, \tilde{\mathbf{y}}_k) \leq v_{J,k}(\mathbf{x}, \mathbf{y}_k) \text{ for } k \in \mathcal{K}. \quad (24)$$

Denote this non-zero sum (Bayesian) game by Γ_K . Note that in game Γ_K a jammer's type has to be associated per a set of channel gains to reflect that the jammer has complete information about channel gains in contrast to the system, which has access to statistical information about them. Of course, in case if both players have access only to statistical information about channel gains there is no need to introduce a such way associated players' types (see, for example, [2]). Here, as examples of using Bayesian game approach in different communication systems we also refer to [10, 13, 15].

Without loss of generality throughout the rest part of the paper we will label the system's strategy $\mathbf{x} = (x, 1 - x)$ by its first component x . Similarly, we will label the k -type jammer's strategy $\mathbf{y}_k = (y_k, 1 - y_k)$ by its first component y_k .

Proposition 1. *In game Γ_K there exists at least one equilibrium.*

The proof can be found in Appendix A.1.

6 Best Response Strategies

In this section we derive in closed form the best response strategies for the system as well as all jammer types.

Proposition 2.(a) For fixed $y_k \in [0, 1], k \in \mathcal{K}$, the best response of system is given as follows:

$$BR_S(y_1, \dots, y_K) \triangleq \begin{cases} \frac{1}{1 + \exp\left(\left(\sum_{k=1}^K \alpha_k D_k y_k - \Theta\right) \delta_w\right)}, & w > 0, \\ \begin{cases} 1, & \sum_{k=1}^K \alpha_k D_k y_k < \Theta, \\ \in [0, 1], & \sum_{k=1}^K \alpha_k D_k y_k = \Theta, \\ 0, & \sum_{k=1}^K \alpha_k D_k y_k > \Theta, \end{cases} & w = 0, \end{cases} \quad (25)$$

where

$$\Theta \triangleq \sum_{k=1}^K (a_k - B_k) \alpha_k, \quad (26)$$

$$\delta_w \triangleq \frac{1-w}{w}, \quad (27)$$

$$D_k \triangleq a_k + b_k - A_k - B_k. \quad (28)$$

(b) For a fixed $x \in [0, 1]$, the best response $y_k = BR_J(x)$ of the k -type jammer is given as follows:

$$BR_{J,k}(x) \triangleq \begin{cases} 0, & x < X_{0,k}, \\ \in [0, 1], & x = X_{0,k}, \\ 1, & x > X_{0,k}, \end{cases} \quad (29)$$

where

$$X_{0,k} \triangleq \frac{b_k - B_k}{D_k}. \quad (30)$$

The proof can be found in Appendix A.2

Remark 1. Note that, by (19), we have that

$$D_k > 0, 1 > X_{0,k} > 0 \text{ and } \delta_w > 0. \quad (31)$$

7 Equilibrium

In this section we derive the equilibrium strategies in closed form and prove their uniqueness. To avoid bulkiness in formulas we assume that $X_{0,k} \neq X_{0,\tilde{k}}$ for $k \neq \tilde{k}$. Then, by (31), without loss of generality we can assume that the channel gain states are arranged in an increasing order by $X_{0,k}$, i.e.,

$$X_{0,1} \triangleq 0 < X_{0,1} < X_{0,2} < \dots < X_{0,K}. \quad (32)$$

Let

$$\varphi_0 \triangleq 0 \text{ and } \varphi_k \triangleq \sum_{i=1}^k \alpha_i D_i \text{ for } k \in \mathcal{K} \quad (33)$$

By Remark 1, φ_k is strictly increasing on k .

Theorem 1. *Let $0 < w < 1$. In game Γ_K the equilibrium is an unique. Moreover, this unique equilibrium $(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)$ is given as follows:*

(a) if

$$\varphi_K - \Theta \leq \ln(1/X_{0,K} - 1)/\delta_w \quad (34)$$

then

$$y_k = 1 \text{ for } k \in \mathcal{K} \quad (35)$$

and

$$x = \frac{1}{1 + \exp((\varphi_K - \Theta)\delta_w)}, \quad (36)$$

(b) if

$$\ln(1/X_{0,1} - 1)/\delta_w \leq \varphi_0 - \Theta \quad (37)$$

then

$$y_k = 0 \text{ for } k \in \mathcal{K} \quad (38)$$

and

$$x = \frac{1}{1 + \exp(-\Theta\delta_w)}, \quad (39)$$

(c) if

$$\varphi_0 - \Theta < \ln(1/X_{0,1} - 1)/\delta_w \text{ and } \ln(1/X_{0,K} - 1)/\delta_w < \varphi_K - \Theta \quad (40)$$

then there exists an unique k_* such that either (41) or (45) hold which are given below, and

(c-i) if

$$\varphi_{k_*-1} - \Theta < \ln(1/X_{k_*,0} - 1)/\delta_w \leq \varphi_{k_*} - \Theta \text{ with } 2 \leq k_* \leq K \quad (41)$$

then

$$y_i = \begin{cases} 1, & i < k_*, \\ \mathcal{Y}, & i = k_*, \\ 0, & i > k_*. \end{cases} \quad (42)$$

and

$$x = X_{0,k_*} \quad (43)$$

where

$$y = \frac{\Theta + \ln(1/X_{k_*,0} - 1)/\delta_w - \varphi_{k_*-1}}{\alpha_{k_*} D_{k_*}}, \quad (44)$$

(c-ii) if

$$\ln(1/X_{k_*+1,0} - 1)/\delta_w \leq \varphi_{k_*} - \Theta < \ln(1/X_{k_*,0} - 1)/\delta_w \text{ with } 1 \leq k_* \leq K - 1 \quad (45)$$

then

$$y_i = \begin{cases} 1, & i \leq k_*, \\ 0, & i > k_* \end{cases} \quad (46)$$

and

$$x = \frac{1}{1 + \exp((\varphi_{k_*} - \Theta) \delta_w)}. \quad (47)$$

The proof can be found in Appendix A.3.

Note that, by (33), $(\varphi_k - \Theta)\delta_w$ is increasing on k . Meanwhile, by (32), $\ln(1/X_{0,k} - 1)$ is decreasing on k . Thus, only one of three conditions (34), (37) or (40) can hold. Also, note that for the boundary case $w = 1$, i.e., if the system is focused on confusing the jammer its equilibrium strategy is $x = 1/2$, where the maximum of entropy is achieved.

8 Discussion of the Results

Let us illustrate the obtained results by an example of the system with $n = 4$ sub-carriers, background noises $\sigma^2 = \sigma_E^2 = 1$, main channel gains $h = (1, 2, 3, 4)$, $h_E = (2, 1, 1, 3)$ and let the total system and jammer power budgets are $\bar{P} = 1$ and $\bar{J} = 5$, respectively. Let the jamming channel gains from the jammer to the communication receiver and radar receiver be $g = (1/d, 1.5/d, 2/d, 0.5/d)$ and $g_E = (2/d, 0.5/d, 1.5/d, 1/d)$, respectively with d being the distance to the receiver from the jammer, where $d \in \{1, 2, 4\}$. Thus, the jamming channel gains can be in $K = 3$ states associated per such a distance. Let $\alpha = (p, (1-p)/2, (1-p)/2)$ with $p \in [0, 1]$ be probabilities that the corresponding channel states occur. Thus, the jamming channel gains might be in state 1 with probability p , meanwhile the rest two states might occur equally probable.

For state 1 we can find power allocation strategies via [1] for the communication and radar modes $\mathbf{P}_{\mathbb{C},1} = (0, 0.22, 0.33, 0.44)$, $\mathbf{J}_{\mathbb{C},1} = (0, 0.44, 1.66, 2.888)$, and $\mathbf{P}_{\mathbb{R},1} = (0.285, 0.142, 0.142, 0.428)$, $\mathbf{J}_{\mathbb{R},1} = (1.571, 0.285, 0.285, 2.857)$,

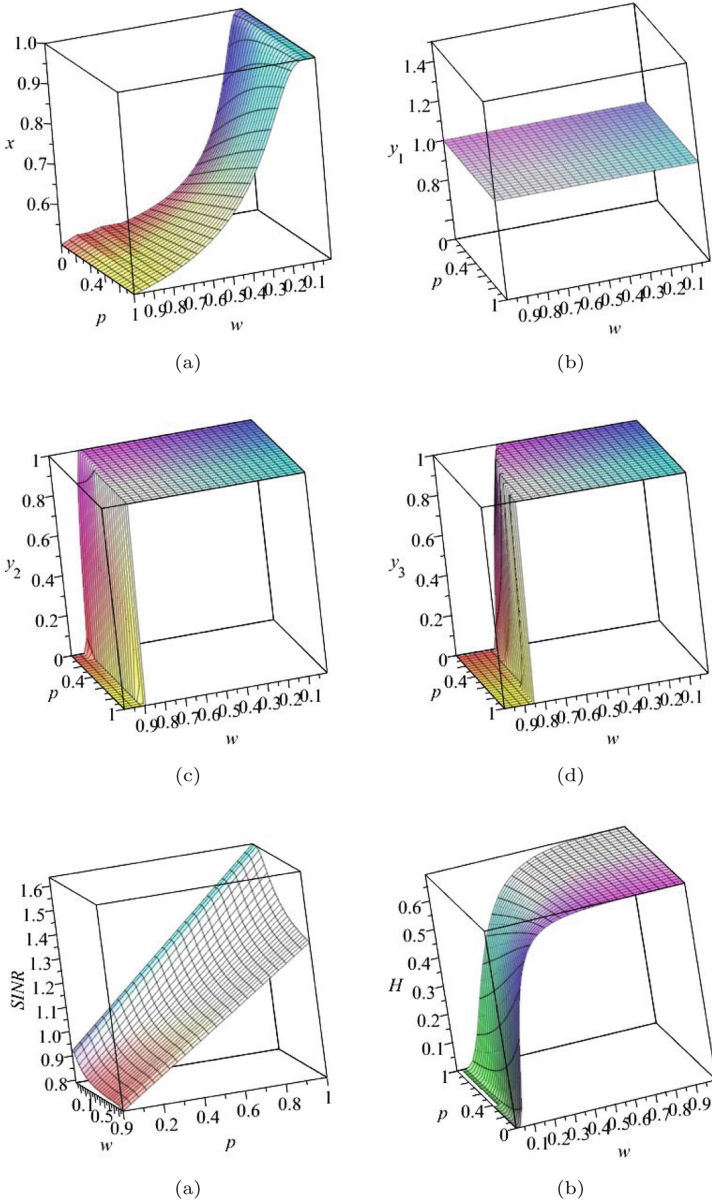


Fig. 3. (a) System strategy x , (b) type-1 jammer strategy y_1 , (c) type-2 jammer strategy y_2 , (d) type-3 jammer strategy y_3 , (e) expected system SINR and (f) entropy $H(x)$ as functions on weighting coefficient w and probability p .

respectively. This leads $A_1 = 1.63$, $a_1 = 1.99$, $B_1 = 0.77$ and $b_1 = 1.05$ as entries of the payoff matrix \mathcal{M}_1 .

For state 2 we can find power allocation strategies in the communication and radar modes $\mathbf{P}_{\mathbb{C},2} = (0, 0.22, 0.33, 0.44)$, $\mathbf{J}_{\mathbb{C},2} = (0, 0.44, 1.66, 2.888)$ and $\mathbf{P}_{\mathbb{R},2} = (0.285, 0.142, 0.142, 0.428)$, $\mathbf{J}_{\mathbb{R},2} = (1.499, 0.499, 0.499, 2.499)$, respectively. This leads to $A_2 = 1.125$, $a_2 = 1.470$, $B_2 = 0.500$ and $b_2 = 0.870$ as entries of the payoff matrix \mathcal{M}_2 .

Finally, for state 3 we can find power allocation strategies in communication and radar modes $\mathbf{P}_{\mathbb{C},3} = (0.10, 0.20, 0.30, 0.40)$, $\mathbf{J}_{\mathbb{C},3} = (0.20, 0.90, 1.60, 2.30)$ and $\mathbf{P}_{\mathbb{R},1} = (0.285, 0.142, 0.142, 0.428)$, $\mathbf{J}_{\mathbb{R},1} = (1.464, 0.607, 0.607, 2.321)$, respectively. This leads $A_3 = 0.714$, $a_3 = 0.896$, $B_3 = 0.291$ and $b_3 = 0.493$ to entries of the payoff matrix \mathcal{M}_3 .

Figure 3 illustrates that the system equilibrium strategy x is continuous with respect to the probability p that state 1 occurs and monotonic decreasing with respect to the weighting coefficient w and tends to $x = 1/2$ for w tending to one, which returns the maximum for entropy, i.e., maximum of unpredictability for the jammer and the minimum for the expected system's SINR. The jammer's type strategies implement boundary values (0 or 1) except maybe the only jammer's type. This reflects the advantage that the jammer has since it has complete information about its channel states in contrast to the system, which has only statistical information. Specifically, this disadvantage makes the system to be more flexible in tuning its strategy in contrast to the jammer since such a task to tune strategy arises for the jammer in the only state.

9 Conclusions

In this paper, in framework of Bayesian game approach, we have modeled a dual purpose communication-radar system that besides the basic objective consisting of two tasks (a) to communicate with a receiver and (b) to track a radar target through the reflections witnessed at the system, also has the secondary objective to achieve the basic objective in a manner that is as unpredictable as possible to the jammer. The entropy associated with the system's strategy to switch between its two tasks of the basic objective has been considered as the metric that reflects the unpredictability of its strategy for the jammer. The uniqueness of the equilibrium is proven, and this reflects the stability of the designed anti-jamming protocol even in the case where the system has less information about the environment than the jammer, as is reflected by knowledge of the underlying jamming fading gains.

A Appendix

A.1 Proof of Proposition 1

By (21), we have that

$$\frac{\partial v_S^2(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)}{\partial x^2} = -\frac{w}{x(1-x)} < 0. \quad (48)$$

Thus, $v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)$ is concave in \mathbf{x} . By (22), we have that $v_{J,k}(\mathbf{x}, \mathbf{y}_k)$ is linear on \mathbf{y}_k , and the result follows from Nash theorem. \square

A.2 Proof of Proposition 2

By (21), we have that

$$\begin{aligned} \frac{\partial v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)}{\partial x} &= (1-w) \left(\sum_{k=1}^K \alpha_k (a_k - B_k) - \sum_{k=1}^K \alpha_k D_k y_k \right) \\ &\quad + w \ln \left(\frac{1-x}{x} \right) \end{aligned} \quad (49)$$

with D_k given by (28). Thus, for a fixed $w \in (0, 1)$ function $\partial v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K) / \partial x$ is decreasing on x from infinity for $x \downarrow 0$ to negative infinity for $x \uparrow 1$. Thus, for a fixed $\mathbf{y} \in [0, 1]$, the best response x is given as the unique root of the following equation:

$$(1-w) \left(\Theta - \sum_{k=1}^K \alpha_k D_k y_k \right) + w \ln \left(\frac{1-x}{x} \right) = 0. \quad (50)$$

Solving this equation by x implies the first row of (25).

For $w = 0$, by (49), we have that

$$v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K) = \left(\Theta - \sum_{k=1}^K \alpha_k D_k y_k \right) x + \sum_{k=1}^K ((b_k - B_k) y_k + B_k). \quad (51)$$

Thus, $v_S(\mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_K)$ is linear in x , and this implies that for a fixed $\mathbf{y}_1, \dots, \mathbf{y}_K$ the best response x is given by the second row of (25)

By (22), we have that

$$v_{J,k}(\mathbf{x}, \mathbf{y}_k) = (B_k - b_k + D_k x) y_k + (B_k - a_k) x - B_k. \quad (52)$$

Thus, $v_{J,k}(\mathbf{x}, \mathbf{y}_k)$ is linear in y_k , and this implies that for a fixed x the best response y_k is given by (29). \square

A.3 Proof of Theorem 1

Let $(x, \mathbf{y}_1, \dots, \mathbf{y}_K)$ be an equilibrium. Then, by (29) and (32), we have that there is a t such that

$$y_i = \begin{cases} 1, & i < t, \\ \in [0, 1], & i = t, \\ 0, & i > t. \end{cases} \quad (53)$$

Let us consider separately three cases for $\mathbf{y}_1, \dots, \mathbf{y}_K$: (a) (35) holds, (b) (38) holds and (c) neither (35) nor (38) hold.

Let (35) hold. This corresponds $t = K + 1$ in (53). Substituting (35) into (25) implies (36). Then, substituting (36) into (29) implies (34), and (a) follows.

Let (38) hold. This corresponds $t = 0$ in (53). Substituting (38) into (25) implies (39). Then, substituting (36) into (29) implies (37), and (b) follows.

(c) Let neither (35) nor (38) hold. Then, by (a) and (b), (40) hold, and, by (29) and (53), two cases arise to consider: (i) $0 < y_t < 1$ and (ii) $y_t = 1$

(i) Let $0 < y_t < 1$. Then, by (29)) and (53), we have that

$$x = X_{0,t}. \tag{54}$$

Substituting (53) into (25), and, then, such obtained x substituting into (54) implies the following equation for y_t :

$$\frac{1}{1 + \exp \left(\left(\sum_{k=1}^{t-1} \alpha_k D_k + \alpha_t D_t y_t - \Theta \right) \delta_w \right)} = X_{t,0}, \tag{55}$$

which, by (33), is equivalent to

$$\varphi_{t-1} + \alpha_t D_t y_t - \Theta = \ln(1/X_{t,0} - 1)/\delta_w. \tag{56}$$

Note that for a fixed t the left side of this equation is increasing on y_t . Thus, (33), this equation has the root y_t in $[0, 1)$ if and only if the following relation holds

$$(\varphi_{t-1} - \Theta)\delta_w \leq \ln(1/X_{t,0} - 1) < (\varphi_t - \Theta)\delta_w. \tag{57}$$

By (33), $(\varphi_t - \Theta)\delta_w$ is increasing on t . Meanwhile, by (32), $\ln(1/X_{0,t} - 1)$ is decreasing on t . Then, since (40) holds, inequalities (57) has the unique solution which we denote by k_* . Finally, solving linear equation (56) on y_t with $t = k_*$ implies (44), and (c-i) follows.

(ii) Let $y_t = 1$. Substituting (53) with such y_t into (25) implies that

$$x = \frac{1}{1 + \exp((\varphi_t - \Theta)\delta_w)}. \tag{58}$$

Then, by (25) and (53) with $y_t = 1$, we have that

$$X_{0,t} \leq \frac{1}{1 + \exp((\varphi_t - \Theta)\delta_w)} < X_{0,t+1}. \tag{59}$$

This inequality is equivalent to

$$\ln(1/X_{t+1,0} - 1) < (\varphi_t - \Theta)\delta_w \leq \ln(1/X_{t,0} - 1). \tag{60}$$

This implies (45), and (c-ii) follows. Finally, existence of the unique k_* given by (41) and (45) follow from the fact that $\ln(1/X_{t,0} - 1)/\delta_w$ is decreasing, meanwhile $\varphi_t - \Theta$ is increasing on t . □

References

1. Altman, E., Avrachenkov, K., GarnaeV, A.: Transmission power control game with SINR as objective function. In: Altman, E., Chaintreau, A. (eds.) NET-COOP 2008. LNCS, vol. 5425, pp. 112–120. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00393-6_14
2. Altman, E., Avrachenkov, K., GarnaeV, A.: Jamming in wireless networks under uncertainty. *Mobile Networks Appl.* **16**, 246–254 (2011)
3. Aubry, A., De Maio, A., Huang, Y., Piezzo, M., Farina, A.: A new radar waveform design algorithm with improved feasibility for spectral coexistence. *IEEE Trans. Aerosp. Electron. Syst.* **51**, 1029–1038 (2015)
4. Bica, M., Huang, K.W., Koivunen, V., Mitra, U.: Mutual information based radar waveform design for joint radar and cellular communication systems. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3671–3675 (2016)
5. Bica, M., Koivunen, V.: Delay estimation method for coexisting radar and wireless communication systems. In: Proceedings of IEEE Radar Conference, pp. 1557–1561 (2017)
6. Fudenberg, D., Tirole, J.: *Game Theory*. MIT Press, Boston, MA (1991)
7. GarnaeV, A., Trappe, W.: An eavesdropping and jamming dilemma with sophisticated players. *ICT Express* (2022). <https://doi.org/10.1016/j.ict.2022.06.002>
8. GarnaeV, A., Trappe, W.: A non-zero sum bandwidth scanning game with a sophisticated adversary. In: Proceedings of 56th Annual Conference on Information Systems and Sciences (CISS), pp. 72–77 (2022)
9. GarnaeV, A., Trappe, W.: A sophisticated anti-eavesdropping strategy. *IEEE Wirel. Commun. Lett.* **11**, 1463–1467 (2022)
10. GarnaeV, A., Trappe, W., Kung, C.-T.: Dependence of optimal monitoring strategy on the application to be protected. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), pp. 1054–1059 (2012)
11. GarnaeV, A., Trappe, W., Petropulu, A.: Bargaining over fair performing dual radar and communication task. In: Proceedings of 50th Asilomar Conference on Signals, Systems, and Computers, pp. 47–51 (2016)
12. Gogineni, S., Rangaswamy, M., Nehorai, A.: Multi-modal OFDM waveform design. In: Proceedings of IEEE Radar Conference, pp. 1–5 (2013)
13. Jean, S., Jabbari, B.: Bayesian game-theoretic modeling of transmit power determination in a self-organizing CDMA wireless network. In: Proceedings of IEEE 60th Vehicular Technology Conference (VTC), vol. 5, pp. 3496–3500 (2004)
14. Li, B., Petropulu, A.P., Trappe, W.: Optimum co-design for spectrum sharing between matrix completion based MIMO radars and a MIMO communication system. *IEEE Trans. Signal Process.* **64**, 4562–4575 (2016)
15. Mohi, M., Movaghar, A., Zadeh, P.M.: A Bayesian game approach for preventing DoS attacks in wireless sensor networks. In: Proceedings of WRI International Conference on Communications and Mobile Computing, vol. 3, pp. 507–511 (2005)
16. Poor, H.V.: *An Introduction to Signal Detection and Estimation*. Springer, New York, NY (1994)
17. Federal Communications Commission (FCC): FCC proposes innovative small cell use in 3.5 GHz band (2012). https://apps.fcc.gov/edocs_public/attachmatch/DOC-317911A1.pdf
18. Turlapaty, A., Jin, Y.: A joint design of transmit waveforms for radar and communications systems in coexistence. In: Proceedings of IEEE Radar Conference, pp. 0315–0319 (2014)

Internet of Things



Power Data Credible Decision-Making Mechanism Based on Federated Learning and Blockchain

Xin Li¹, Fangjian Shang¹, Yanli Yao¹, and Tianren Zheng²(✉)

¹ State Grid Jibei Information and Telecommunication Company,
Beijing 100053, People's Republic of China

{li.xin.q, shang.fangjian, yao.yanli}@jibei.sgcc.com.cn

² Beijing University of Posts and
Telecommunications, Beijing 100876, People's Republic of China
zhtr@bupt.edu.cn

Abstract. In modern power systems, it is an important issue to process and analyze power big data and perform reliable decision-making analysis. In response to this problem, this paper proposes a distributed computing architecture for power data based on a consortium chain, which realizes distributed and trusted shared training computing for power data while taking into account the privacy protection of the original data. To solve the problem of sample imbalance, this paper proposes a data balancing method combining SMOTE algorithm and the k-means algorithm. This paper also proposes an LSTM neural network load forecasting method based on federated learning and proves that it has higher accuracy and applicability than traditional methods through examples.

Keywords: Power data · Federated learning · Blockchain · LSTM

1 Introduction

Modern power systems have built a series of advanced intelligent monitoring infrastructures. These devices can generate a large amount of data. Processing and analyzing these data can get a lot of effective information, to understand the operating status of power equipment and the power consumption of users. Through the analysis and research of historical data, the future electricity consumption data can be predicted, to make a series of decisions on the change of the operating state of power equipment. Smart grids are designed to save energy, reduce losses, and enhance grid reliability. This puts forward higher requirements for credible decision-making in the power system. How to process and analyze power big data and perform reliable decision-making analysis has become important for research problem. In the field of power systems, making full use of the value of existing data research, fusing machine learning models and artificial neural networks can provide accurate forecasts for power loads, and provide reference

and decision-making guidance for power generation, power sales, and power consumption. Real-time high-accuracy load forecasting will promote the development of credible decision-making and further promote the development of smart grids.

Blockchain technology has great advantages in terms of security. The characteristics of high data redundancy and resistance to tampering are very consistent with the strict requirements for data security in the power system. Each node in the blockchain network stores complete data, and the consistency of the data is ensured through a consensus mechanism. Blockchain technology was first applied in the field of digital encryption currency, and its high security based on cryptographic principles has been widely recognized, and then it has been introduced into more and more new application scenarios. In the energy field, blockchain technology has been tried to be applied in fields such as electricity trading, carbon emission rights trading, and more and more other businesses.

In modern power systems, in many cases, the data set cannot be centralized or shared between the two parties' data, and cannot be used for data mining and analysis, so it faces the problem of data islands. If the power data is centralized for analysis and utilization, it may also involve laws and regulations, user privacy, and data security issues. To solve this problem, this article introduces a federated learning method. The federated learning can ensure that the data can be trained locally without the data. This can protect the data security and user privacy of the edge nodes to a certain extent, and reduce the security risk in the data transmission process.

2 Related Work

The combination of blockchain and energy Internet of Things can promote the marketization of energy and the intelligentization of the grid system, which has huge development potential. Literature [1] analyzed the application applicability of blockchain in multi-energy systems and the information interconnection problems brought about by heterogeneous blockchains and proposed the construction of a multi-energy system transaction system based on heterogeneous blockchain technology. Necessity and method. Literature [2] proposes a blockchain system suitable for distributed photovoltaic microgrids, with the help of digital currency, photovoltaic transactions can be carried out without being monitored.

Regarding federated learning, in 2019, Google used the federated learning platform to train LSTM neural networks to learn vocabulary outside the vocabulary [3]. Zhao Tao et al. [4] proposed a federal learning aviation travel prediction method oriented to data privacy protection, which greatly improved the accuracy and reliability of aviation travel prediction.

Regarding power load forecasting, there have been decades of research in academia, and many effective forecasting methods have been proposed. In the 1970s and 1980s, researchers often used regression analysis, time series, and other forecasting methods. These methods are generally called classic forecasting methods. Literature [5] aimed at regression-based prediction models, using modern computing power to customize models for two to three years of hourly data to maximize prediction accuracy. Literature [6] expresses load forecasting as a functional time series problem and uses the function wavelet kernel method to predict the load curve of the clusters divided into groups.

With the rapid development of the energy Internet of Things, the amount of load data is increasing, and the demand for accuracy of power load forecasting is also increasing. In recent years, deep learning has made great progress. Therefore, the current research on short-term load forecasting mainly focus on neural network. Literature [7] trains a neural network on multiple data sets generated by random sampling and replacement and then averages the results to reduce load forecast errors. In [8], to overcome the short-term load forecasting (IGRA-BA-BP) based on improved gray correlation and bat optimized neural network, to overcome the disadvantages of back-propagation (BP) neural network with poor generalization ability and easy to fall into local optimum. Method, the forecasting effect has been improved. Literature [9] established a cloud power load forecasting platform, based on the Markov chain to reduce forecasting errors from the bottom up.

3 Power Data Distributed Computing Architecture Based on Alliance Chain

As shown in Fig. 1, based on the existing consortium chain architecture of the power grid system, a distributed computing model and framework based on the consortium chain can be constructed to realize distributed and trusted shared training and computing of power data while taking into account the protection of original data privacy. The architecture consists of the data collection layer, the national network chain layer, the federation learning layer, and the algorithm application layer. The bottom layer is the data collection layer, which is mainly responsible for the collection of raw power data in various power parks, and then the collected data is Uploaded to the edge servers of the power grid, these edge servers constitute the working nodes of the alliance chain network of the upper state grid chain layer. Finally, through the design of smart contracts, each working node is organized to complete the uppermost federal learning application based on power data. Decision-making and analysis tasks are made by the level. In the process of specific algorithm application, taking into account the characteristics of power data sharing, this project introduces a data balance method combining SMOTE algorithm and K-means clustering algorithm into the data layer of the joint decision-making system to balance positive and negative samples. Use LSTM neural network to predict and analyze power load data to complete related decision-making tasks.

4 Data Credible Decision-Making Mechanism

4.1 Date Balance Method Combining SMOTE Algorithm and K-means Algorithm

Due to the robustness of the power system in practice, the system can often recover to a steady state by itself after being disturbed, and the probability of instability is relatively low, which brings the problem of sample imbalance to the method based on deep learning.

A typical oversampling method to solve the problem of data imbalance is the SMOTE algorithm, which aims to make up for the shortcomings of a small number of random oversampling. Random oversampling of the minority samples make the minority samples

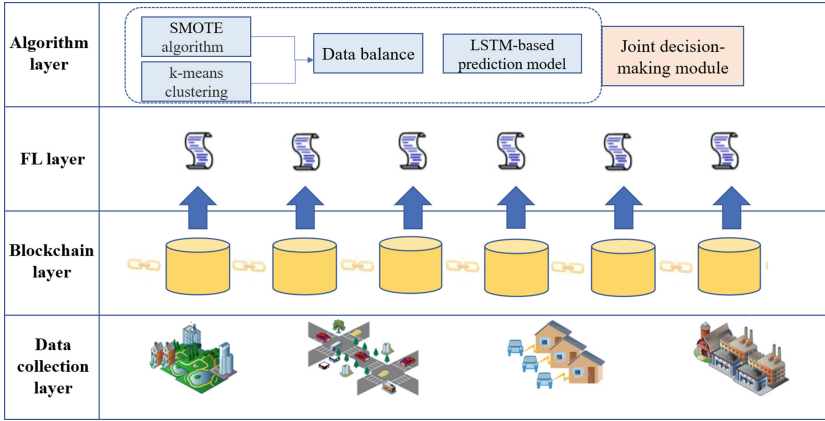


Fig. 1. Power data distributed computing architecture based on alliance chain

more recognizable, because the oversampling process is actually copying the samples. This copying will make the decision-making decisions oToes more and more rigorous and specific, leading to Classification overfitting. In order to solve this problem, SMOTE introduce synthetic data points and line segments adjacent to any 1 or all of k nearest neighbors of a minority class in the characteristic space. If (x_1, x_2) is an instance of a minority class, and if its nearest neighbor is selected as (x_1', x_2') , then the data (X_1, X_2) is synthesized, namely:

$$(X_1, X_2) = (x_1, x_2) + \text{rand}(0, 1) \times \Delta \tag{4.1}$$

Among them, $\Delta = \{(x_1' - x_1), (x_2' - x_2)\}$, $\text{rand}(0, 1)$ is a random number between 0 and 1. This technique broadens the decision-making area by generating artificial samples, because the samples added to the data set are located in the vicinity of the original samples, rather than the samples themselves. Compared with random oversampling with replacement, the decision area is more general.

The traditional SMOTE algorithm still has some problems. For example, the use of the SMOTE algorithm may blur the positive and negative class boundaries of the data set, which will increase the difficulty of training the classification model, and the SMOTE algorithm has a certain degree of blindness when processing data. In order to solve these problems. A clustering algorithm can be introduced to cluster the minority classes before oversampling, and then sample the clusters after clustering. This paper introduces the combination of K-means clustering algorithm and SMOTE algorithm. The K-means clustering algorithm takes the distance between the sample point and the cluster center as the optimization goal. According to the core idea of the clustering algorithm, the algorithm will maximize the similarity of the elements in each cluster as much as possible, and the difference between clusters The similarity is minimized. The K-means algorithm selects the desired clusters, through continuous iteration and recalculation of the cluster centers to minimize the variance within the entire cluster, and obtains relatively compact and independent clusters as the final goal of the algorithm.

Use the function method to obtain the extreme value, adjust the threshold of the number of iterations to obtain the best clustering effect.

The specific steps of the K-means algorithm are as follows:

STEP1 For data set D , randomly select k initial cluster centroid points as $\mu_1, \mu_2, \dots, \mu_k \in D^n$.

STEP2 For the data except for the cluster center, calculate the Euclidean distance between them and $\mu_i (i = 1, 2, \dots, k)$ one by one, and group the data closest to $\mu_i (i = 1, 2, \dots, k)$ together so that all the data are divided into k categories. Can use formula (4.2) to classify

$$c^{(i)} = \arg \min_j \|x^{(i)} - \mu_j\|^2 \quad (4.2)$$

STEP3 Calculate the mean value of the data in various clusters and set the obtained mean value to the center of the new cluster. Then calculate the sum of the Euclidean distance from each data point in this cluster to the center of the cluster.

$$J(c, \mu) = \sum_{i=1}^m \|x^{(i)} - \mu_{c(i)}\|^2 \quad (4.3)$$

STEP4 Repeat the second and third steps. If the sum of Euclidean distance J does not change, output the clustering result.

Suppose a given data set D , where the majority class sample set is D_{max} , the minority class set is D_{min} , p is the number of minority class samples, and the sampling magnification is n . The specific description of the data balance method combining SMOTE algorithm and K-means algorithm is as follows:

STEP1 According to the specified parameter t of the given data set, randomly select t sample points in the data set (the selected sample points must belong to the minority class), and divide the minority class samples into t clusters $T_i (i = 1, 2, \dots, t)$.

STEP2 Standardize the data in the data set and scale the data according to a certain ratio. After processing, the values of all attributes in the data set are in the interval $(0,1)$.

STEP3 judges the number of most sample points in each cluster, that is $|T_i \cap D_{max}| = q$, if $q = |T_i|$, the cluster belongs to the noise cluster and the set is N' ; if $q > |T_i|/2$, the cluster belongs to the boundary cluster and the set is B' ; if $0 < q < |T_i|/2$, the cluster belongs to the safe cluster and the set is S' .

STEP4 removes the majority samples in the boundary clusters and only retains the minority samples. With each boundary cluster as the unit and the respective cluster center as the core, new sample points are generated according to formula (4.4).

$$Y_{new} = c_i + RAND(0, 1) \times (y_j - c_i) \quad (4.4)$$

STEP5 synthesizes the generated sample points into the minority samples to determine whether the overall data set is in balance. If it does not reach the balance, iterate 3 and 4 processes until the data set is close to balance, then the algorithm ends.

STEP6 reverse-standardizes the obtained data set, and converts the values in the interval $(0,1)$ into the indicators of the original data according to the ratio of each attribute in step 2, and restores the attributes of the data in the original data set.

4.2 Load Forecasting Module Based on LSTM Neural Network

In order to deal with the problem of gradient disappearance that often occurs in RNN, Hochreiter et al. proposed a long short term memory (LSTM) neural network based on RNN. This is an effective non-linear cyclic neural network, which can take care of the time series and non-linear relationship of the data, and has emerged in load forecasting. This paper proposes a load forecasting method based on LSTM.

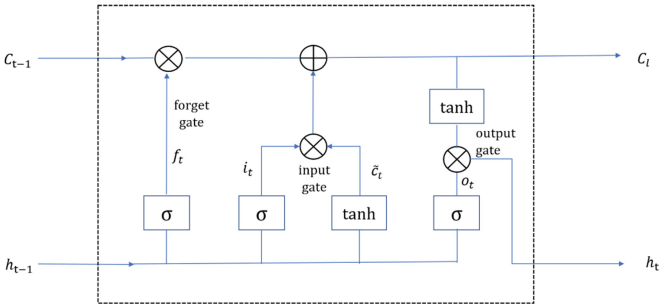


Fig. 2. Schematic diagram of LSTM structure

As shown in Fig. 2, in order to prevent the gradient extinction problem similar to RNN, LSTM adopts a gate structure to strengthen the information transmission between each neuron. It consists of three gate structures: input gate, output gate, and forget gate. Responsible for controlling the input, output and historical dependence of cells, and work together to realize load forecasting of power data. The specific operation process of LSTM is as follows

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{4.5}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{4.6}$$

$$\tilde{c}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{4.7}$$

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \tag{4.8}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{4.9}$$

$$h_t = \sigma_t \tanh c_t \tag{4.10}$$

Equation (4.5) controls the information that the current neuron forgets in the previous neuron, which is realized by the Sigmoid layer of the forget gate. Operate by reading the output h_{t-1} of the previous neuron and the input x_t of the current neuron to output a value in the interval [0,1], where 1 represents the complete memory of the previous neuron

state, and 0 represents the previous neuron state. The neuron state is completely forgotten, and then multiplied by the previous neuron state c_{t-1} . In this way, the long-term memory of the neural network is guaranteed.

Equations (4.6) and (4.7) jointly control the input of neurons. Equation (4.6) is realized by the sigmoid layer of the input gate. The input gate reads the output h_{t-1} of the previous neuron and the input x_t of the neuron outputs a value i_t in the interval $[0,1]$. Equation (4.7) generates a candidate value vector \tilde{c}_t through the tanh layer, and then obtains the neuron state c_t through (4.8).

Equation (4.10) gets the final output. After calculating the information retained by the previous neuron and the information of the neuron, the sigmoid layer of the output gate will get the final output. The neuron state c_t is processed through the tanh layer, and then multiplied by the output gate σ_t to obtain the final output h_t .

For the current predicted moment t , the power data from moment $t - n$ to moment $t - 1$ is used as model input X , and the power data from moment t is used as output y , as shown below

$$X_t = [d_{t-n}, d_{t-n+1}, \dots, d_{t-1}] \tag{4.11}$$

$$y_t = d_t \tag{4.12}$$

where, d represents power data.

Federated learning solves the problem of data silos. The model training and sharing through federated learning mainly consists of two parts: organization model training and cloud model aggregation. The training objectives of cloud model and each node mechanism model can be expressed as:

$$\arg \min_{\omega, b} L = \sum_{i=1}^n l(y_i, f_S(x_i)) \tag{4.13}$$

$$\arg \min_{\omega^j, b^j} L_j = \sum_{i=1}^{n^j} l(y_i^j, f_j(x_i^j)) \tag{4.14}$$

wherein, ω and b represent the training parameters: weight and deviation, L represents the loss function, (x_i, y_i) and (x_i^j, y_i^j) represent the global power data and the power data of the JTH node mechanism, and n represents the data set size.

The algorithm flow of federal learning and training LSTM neural network model is as follows:

Algorithm 1 Load prediction model training process

Input: common data set D_0 , different mechanism data set $\{Q_1, Q_2, \dots, Q_j\}$ **Output:** Final model F1: LSTM model f_s is trained by using D_0 in cloud server

2: FOR round = 1, 2, ..., r DO

3: Send the model f_s to all mechanism nodes4: Each node uses local data D_j to train its own local model f_j and upload it to the cloud server

5: The cloud server aggregates all model parameters and updates the global model

6: END FOR

5 Performance Evaluation

Footnotes should be avoided whenever possible. If required they should be used only for brief notes that do not fit conveniently into the text.

5.1 Data Set

The power load data of Slovenia from 2020-11-01 to 2021-11-01 was downloaded from the ENTSO website as an experimental data set. The data set consists of power load sampling every 1 h, totaling 8784 (24×366) pieces of power data (Fig. 3).

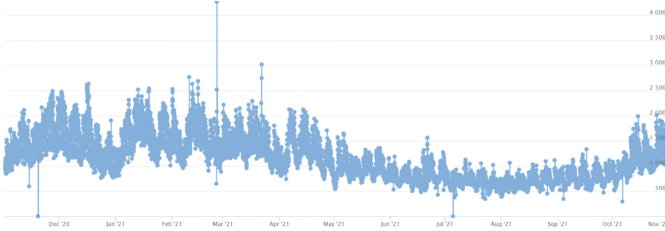


Fig. 3. Data set

5.2 Model Structure Design

The experimental data is divided into two parts. The first 51 weeks are used as the data training set, and the load data on October 30 and 31 are used as the test set to test the effect of the model in comparison with the predicted results. Input the training data into RNN and LSTM models for training, the number of hidden layer neurons (num_units) is designed to be 128, the training data batch (batch_size) is 16, the number of data in each batch (window_size) is 400, and the learning rate (AdamOptimizer) is 0.001, and the number of iterations (train_steps) is 3000.

5.3 Result Analysis

The average percentage error is used to compare the training effects of the two methods. The average percentage error is defined as (Figs. 4 and 5)

$$E_{MAPE} = 100 \frac{\sum_{i=1}^n \left| \frac{L_i - L_i'}{L_i} \right|}{n} \tag{5-1}$$

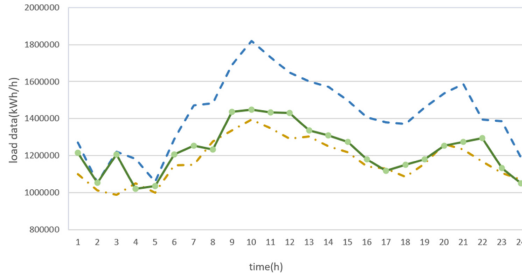


Fig. 4. Comparison of 2021.10.30 and real data

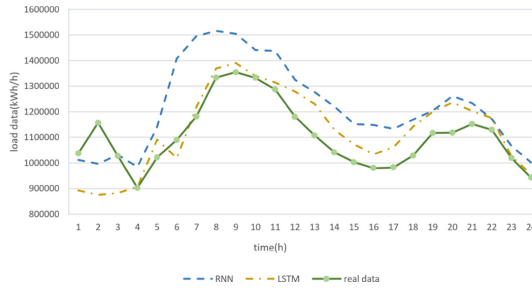


Fig. 5. Comparison of 2021.10.31 with real data

where L_i and L_i' are the real value and the predicted value respectively.

Calculate E_{MAPE} for each method, as shown in Table 1.

Table 1. Comparison of E_{MAPE} between the two methods

Method\Data	10.30	10.31
RNN	15.88	11.92
LSTM	5.12	6.90

It can be clearly seen from the above table that the average percentage error of the prediction results of the LSTM neural network is much better than that of the RNN

network. This is because the RNN model can only learn the diurnal variation of power data, and the LSTM model can learn better Daily and weekly changes. This proves that the LSTM method is better than the traditional method in the use effect.

Acknowledgments. This paper is supported by the science and technology project of State Grid Jibei Information & Telecommunication Company “Trusted Sharing Technology and Application of Winter Olympic Power Data Based on Blockchain” (52018E20008J).

References

1. Li, B., et al.: Transaction system and key technologies of multi-energy system based on heterogeneous blockchain. *Autom. Electr. Power Syst.* **42**(4), 183–193 (2018)
2. Johnson, L., Isam, A., Gogerty, N., Zitoli, J.: Connecting the Blockchain to the Sun to Save the Planet. Available at SSRN 2702639 (2015)
3. Beaufays, F.S., Chen, M., Mathews, R., Ouyang, T.: Federated learning of out-of-vocabulary words (2019)
4. Chen, T., Guo, R., Liu, Z.: Research on aviation application model of federated learning algorithm for big data privacy protection. *Inf. Secur. Commun. Secrecy* **09**, 75–84 (2020)
5. Wang, P., Liu, B., Hong, T.: Electric load forecasting with recency effect: a big data approach. *Int. J. Forecast.* **32**(3), 585–597 (2016)
6. Chaouch, M.: Clustering-based improvement of nonparametric functional time series forecasting: application to intra-day household-level load curves. *IEEE Trans. Smart Grid* **5**(1), 411–419 (2013)
7. Khwaja, A.S., Naeem, M., Anpalagan, A., Venetsanopoulos, A., Venkatesh, B.: Improved short-term load forecasting using bagged neural networks. *Electr. Power Syst. Res.* **125**, 109–115 (2015)
8. Wu, Y., Lei, J., Bao, L., Li, C.: Short-term load forecasting based on improved grey relational analysis and neural network optimized by Bat algorithm. *Autom. Electr. Power Syst.* **42**(20), 67–72 (2018)
9. Gong, X., Cardenas-Barrera, J.L., Castillo-Guerra, E., Cao, B., Saleh, S.A., Chang, L.: Bottom-up load forecasting with Markov-based error reduction method for aggregated domestic electric water heaters. *IEEE Trans. Ind. Appl.* **55**(6), 6401–6413 (2019)



A Service Protection Mechanism of Deterministic Networking Based on Segment Routing

Tianchi Li^(✉) and Yueping Cai

Chongqing University, Chongqing, China
{20144009, caiyueping}@cqu.edu.cn

Abstract. Deterministic Networking (DetNet) provides guaranteed packet transport services of ultra-low packet loss and bounded delay for the critical traffic in real-time applications such as the industrial control and the power grid. DetNet guarantees reliable packet transmissions by forwarding replicated packets on redundant paths in parallel. This service protection mechanism of DetNet is Packet Replication Elimination and Ordering Functions (PREOF). However, how to obtain the redundant paths and implement the packet replication and elimination functions of the PREOF remains to be a great challenge. This paper proposes an improved PREOF mechanism based on Segment Routing (SR-PREOF). It designs an edge-disjoint path-pair routing algorithm based on the improved Link Pruning method (LP-EDJPP). The proposed SR-PREOF implements the scheme with the SR technology. Network simulation results show that the proposed SR-PREOF effectively improves the packet reception rate and reduces the end-to-end worst-case latency bound while achieving the comparable path reliability performance compared with the traditional PREOF. The packet reception rate of the SR-PREOF increases by 5.6% and the end-to-end worst-case latency bound decreases by 10.89% compared to the PREOF when the offered load is 0.7.

Keywords: Deterministic Networking · Service protection · Segment routing · Routing algorithm

1 Introduction

The rapid development of Industrial Internet of Things brings great changes to the production and life style of human beings, but it presents more strict requirements for network services. Industries such as the industrial automation, control system and power grid have strong demands for *deterministic* network services with bounded delay and ultra-low packet loss [1]. However, the traditional IP network provides users with *best effort* service, which is difficult to meet the strict Quality of Service (QoS) requirements of deterministic traffic on the delay, jitter and packet loss rate. The emergence of Deterministic Networking (DetNet) [2] provides a new solution for the traffic demands of such industries.

DetNet is a network technology that provides a promised Service Level Agreement (SLA) guarantee for the time-sensitive traffic. It enables the guaranteed QoS services through the mechanisms of the explicit routes, resource allocation and service protection. Moreover, it guarantees the performance of the time-sensitive traffic in terms of the delay jitter, packet loss and end-to-end (E2E) bounded latency.

Three key enabling technologies of DetNet include: 1) **Explicit routes**: Provide explicit paths for the DetNet flows that meet the user's SLA requirements. 2) **Resource allocation**: Book data plane resources for the DetNet flows according to the allocated paths. 3) **Service protection**: Replicate the DetNet flows and transmit the copies in parallel over different paths. The service protection is designed to reduce or eliminate packet loss due to equipment or link failures. The packet loss can be greatly reduced by spreading the data over multiple disjoint paths for transmissions. Traditional network protection methods include $1 + 1$ linear protection [3], network coding [4], PROEF [5] etc. and all of them consume extra network bandwidth resources. The Packet Replication Elimination and Ordering Functions (PREOF) mechanism is proposed in [5] to solve the service protection problems of DetNet. Nevertheless, there are no mature routing algorithms and implementation details of the PREOF. In addition, the efficiency of the method needs to be improved.

Segment Routing (SR) [6] is based on Multi-Protocol Label Switching (MPLS) [7], which has the features of source routing and header instruction programmability. It supports explicit routes on IPv6 or MPLS data plane, which can provide the powerful traffic scheduling capability. This paper proposes an improved PROEF mechanism based on Segment Routing (SR-PREOF). It designs an edge-disjoint path-pair routing algorithm based on the improved Link Pruning [8] method (LP-EDJPP). It can compute the appropriate replication node and elimination node as well as the edge-disjoint path-pair to transmit deterministic traffic flows. The proposed SR-PREOF implements the service protection with the SR technology. Network simulation results show that the proposed SR-PREOF effectively improves the packet reception rate and reduces the E2E worst-case latency bound while achieving the comparable path reliability performance compared with the traditional PROEF. The packet reception rate of the SR-PREOF increases by 5.6% and the E2E worst-case latency bound decreases by 10.89% compared to the PROEF when the offered load is 0.7.

The contributions of this paper are as follows:

- 1) An edge-disjoint path-pair routing algorithm based on the improved Link Pruning method (LP-EDJPP) is designed considering about the constraints of network topological connections and delays.
- 2) An improved PROEF mechanism based on Segment Routing (SR-PREOF) is proposed. It utilizes SR technology to implement the scheme for DetNet.

2 Related Work

The whole process of the service protection for DetNet can be described as following steps [2]. First, provide *sequencing information* to the packets of a DetNet Compound Flow (DCF) at the source node. This can be done by adding a *sequence number* or

timestamp as part of the defined DCF. Second, replicate the packets into multiple DetNet Member Flows (DMFs). It needs an appropriate replication node. Third, send the packets on multiple different paths to the destination node. In this way, when some links of one path fail, the transmissions of packets on the other paths are not affected. Finally, eliminate duplicated packets and order them at an appropriate *elimination node*. However, there is a challenge on finding suitable disjoint paths for packets in DetNet.

Link Pruning [8] is one of the traditional disjoint paths routing algorithms. The steps of Link Pruning include: 1) Find one path with the minimum cost from the given topology diagram. 2) Delete all the links on the path in the topology diagram. 3) Find another path with the minimum cost from the pruned topology diagram. However, this method cannot guarantee the consistency of QoS constraints and even cannot find two disjoint paths sometimes. Besides, other people did many works on disjoint paths routing algorithms. Huong T.T. et al. [9] proposed a global load-balanced routing scheme, which could take advantage of global view of the SDN controller to make a global policy for routing and load balancing. Weiner J. et al. [10] proposed a new method to solve the Maximum Edge Disjoint Paths (MEDP) problem. The proposed method was a heuristic algorithm that built a hybridization of Lagrangian Relaxation and Particle Swarm Optimization. Atallah A. A. et al. [11] designed a disjoint multi-path QoS routing algorithm for E2E networks. Jonatan K. et al. [12] focused on the joint routing and scheduling problem of DetNet. This paper proved that the disjoint routing problem of DetNet is NP-hard and the heuristic algorithm is needed to solve it. Aubry F. et al. [13] proved that SR technology is an approach to realize robustly disjoint paths in Internet Service Providers (ISPs). This paper introduced the theory and model for the robustly disjoint paths. It exploited SR's ability to implicitly specify backup paths. Aubry F. et al. [14] proposed an efficient algorithm that computed K-segmentable disjoint paths with similar latencies. It was possible to provide an $I + I$ protection service by using segmentable disjoint paths in an IPv6 network that supports Segment Routing architecture.

The Packet Replication, Elimination and Ordering Functions (PREOF) mechanism was proposed in [5] as a typical solution for the service protection in DetNet. PREOF includes three parts: 1) **In-order delivery**. Packets delivered out-of-order will increase the amount of buffering needed at the destination but also the jitter. The Packet Ordering Function (POF) uses the *sequencing information* to re-order a DetNet flow's packets that are received out-of-order. 2) **Packet replication**. The Packet Replication Function (PRF) replicates these packets into multiple DMFs, and typically sends them along multiple different paths to the destination. 3) **Packet elimination**. The Packet Elimination Function (PEF) eliminates duplicate packets of a DetNet flow based on the *sequencing information* and the history of received packets. In general, PREOF is applied to the DetNet service sub-layer for the packet processing at edge nodes, relay nodes and end systems of DetNet. In the simplest case, each packet is replicated at the source node and transmitted over two disjoint paths to the same destination. However, there are still the following problems to be solved. 1) A method to choose the appropriate replication node and elimination node according to the network topology. 2) An algorithm to compute disjoint path-pair as close in costs as possible. It will help the in-order delivery and lead to less E2E jitter and delay. 3) A programmable implementation deployed on the DetNet data plane.

In this paper, the traditional PREOF is extended and improved. The proposed SR-PREOF mechanism utilizes the LP-EDJPP algorithm to find the appropriate nodes and paths, and implements the service protection in DetNet over IPv6 data plane with the SR technology.

3 Service Protection Mechanism SR-PREOF

The improved PREOF mechanism based on Segment Routing (SR-PREOF) mainly includes three parts: path calculation, SR policy programming and data plane forwarding. Figure 1 shows the schematic diagram of the proposed mechanism. This section shows how the SR-PREOF mechanism can operate within the E2E DetNet domain. The proposed LP-EDJPP algorithm can be implemented in the centralized Path Calculation Module (PCM). In the proposed mechanism, the disjoint path-pair computed by the PCM can be implemented as a sequence of Segments. The SR Policy Programming Module (SRPPM) can encode the path-pair information into the Segment Routing Header (SRH). And it informs the data plane to forward packets according to the SR policy. The Data Plane Forwarding Module (DPFM) forwards packets based on the delivered SR policy.

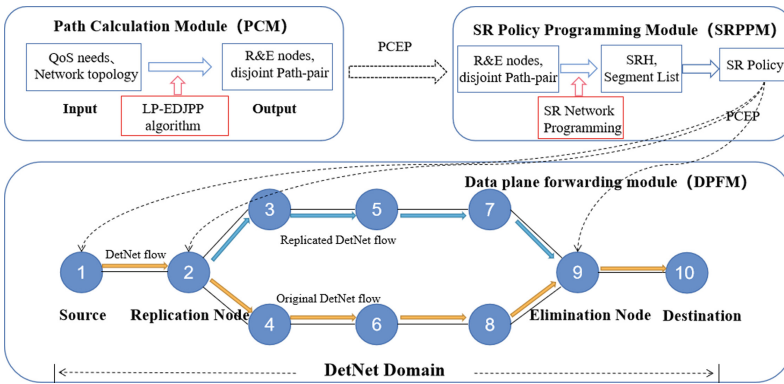


Fig. 1. Schematic diagram of the Deterministic Networking service protection mechanism: SR-PREOF

3.1 Path Calculation

The problem of the path calculation formulated in this paper can be described as follows. Network topology G , source node S and destination node D are given. Computing two disjoint paths between the two nodes and the two paths satisfy a certain objective function. The objective function here is MinSum function and the problem is called MinSum-Balanced 2DP Problem [15]. In this problem, the two disjoint paths we get have the smallest sum of weights and the absolute value difference between the weights of the two paths is the smallest. The problem was proved NP-Completeness in undirected graphs. Therefore, we propose a heuristic algorithm based on the shortest paths approach.

The path calculation module adopts the edge-disjoint path-pair routing algorithm based on the improved Link Pruning (LP-EDJPP) to select edge-disjoint path-pair, the replication node and elimination node for the service protection. The traditional Link Pruning algorithm has following problems: 1) The deletion of links may cause that two disjoint paths cannot be found. 2) The number of links between two disjoint paths may be too different to ensure the consistency of QoS performance constraints. In addition, it may result in out-of-order packet arrivals. To overcome the above problems, we design the LP-EDJPP algorithm. The aim of the LP-EDJPP algorithm is to find the equal-cost edge-disjoint path pairs. The proposed LP-EDJPP algorithm is an approximation algorithm, and the complexity of it is $O(N^2)$. It can get a quality-guaranteed approximate solution in polynomial time.

In this paper, the cost is the latency. In addition, the designed algorithm only considers the edge disjoint (link disjoint) but does not consider the node disjoint. Finding path-pair with completely disjoint nodes and edges will cause excessive delay and bandwidth consumption. The specific process is as follows.

Network Model. The network topology can be represented by an undirected weighted graph $G = \{V, E, C\}$, where V is a set of nodes, E is a set of edges and C is the edge weight set which is a function from E to R^+ . This function corresponds to the latencies configured on the edges. Assume that $|V| = N$, $|E| = M$, then the node set $V = \{v_1, v_2, \dots, v_n\}$, the edge set $E = \{e_1, e_2, \dots, e_m\}$. The edge weight set $C = \{w_1, w_2, \dots, w_m\}$. (u, v) represents an edge between node u to node v , where $u, v \in V$. $w = (u, v)$ represents the weight of the edge.

A path P is a sequence (v_1, v_2, \dots, v_h) such that $(v_i, v_{i+1}) \in E$ for all i and $v_i \neq v_j$ for $i \neq j$. $P(s, d, G)$ represents the shortest path of node pair (s, d) . Path-pair X-Y disjoint means $E(X) \cap E(Y) = \emptyset$.

Shortest Path Calculation. Dijkstra [16] algorithm is the classical shortest path algorithm, which can find the shortest path of any two nodes in the graph. The time complexity is $O(N^2)$. We can compute the shortest path $P = (v_1, v_2, \dots, v_p)$. Assume that the network topology $G = \{V, E, C\}$ is given as shown in Fig. 2. The shortest path $P(1, 9, G) = 1-3-6-8-9$ between the node pair $(1, 9)$ can be easily calculated according to the Dijkstra algorithm.

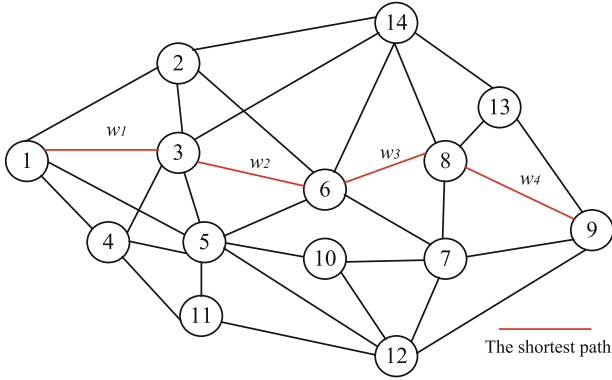


Fig. 2. Schematic diagram of network topology G and the shortest path

Edge-Disjoint Path Finding. Set $w(v_i, v_{i+1})$ on path P to w^* , generate a new network topology $G' = (V, E, C')$, then use shortest path calculation algorithm to compute another path P' . In this case, as shown in Fig. 3, set the weight of each edge on P ($1, 9, G$) to w^* . Note that w^* is a constant, where $w^* = \text{Sum } C(P)$. Run the Dijkstra algorithm based on the new network topology, calculate another path P' ($1, 9, G'$) = $1-2-6-7-9$.

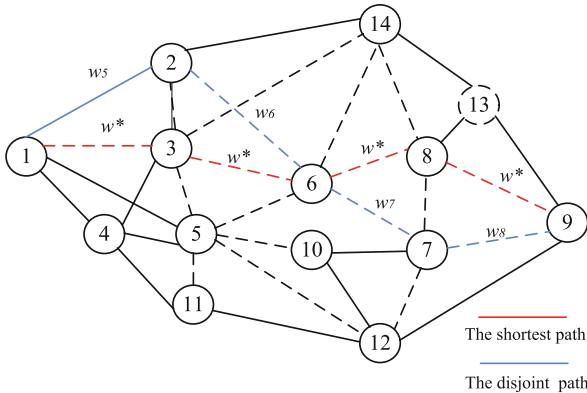


Fig. 3. Schematic diagram of network topology G' and the edge-disjoint path

Obtain the Edge-Disjoint Path-Pair. If the two paths are node-disjoint, then the path P and P' are what we need. Most of the time, we get two paths that have intersection nodes. In this case, we get a new graph G^* as shown in Fig. 4. In order to get the equal-cost edge-disjoint path-pair, we need a further optimization. Assume the edge state is $0-1$ variables. 0 means unused, 1 means used. The path-pair is $X-Y$, the edge weight set is W . The optimizing goal is to minimize the latency gap between the path X and Y , as shown in Eq. (1).

$$\text{Min}((X - Y) \cdot W) \tag{1}$$

$$s.t. X, Y \in G^* \quad (2)$$

We formulate it to an Integer Linear Programming problem, which can be easily solved. In this case, $W = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\}$, and the edge state is: path $X = [1, 1, 1, 1, 0, 0, 0, 0]$, path $Y = [0, 0, 0, 0, 1, 1, 1, 1]$. Assume that $|(w_1 + w_2 + w_7 + w_8) - (w_3 + w_4 + w_5 + w_6)|$ is the minimum, then we get the edge-disjoint path pair X-Y, where $X = 1-3-6-7-9$, $Y = 1-2-6-8-9$.

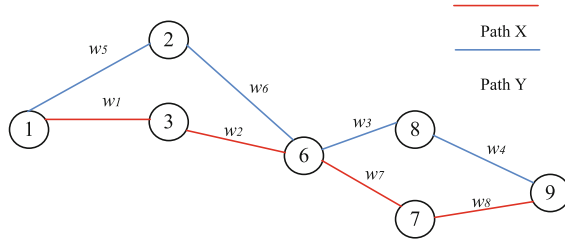


Fig. 4. Schematic diagram of network topology G^* and the edge-disjoint path-pair X-Y

The pseudo-code of LP-EDJPP is as follows:

Algorithm 1: Edge-disjoint path-pair routing algorithm based on Link Pruning

- 1) **Input** G : network topology
 - 2) s, d : source and destination node
 - 3) **Output** X, Y : edge-disjoint path-pair
 - 4) $V = \{r, e\}$: replication and elimination node
 - 5) **Start**
 - 6) **Dijkstra** $(s, d, G) \rightarrow P(s, d, G)$; // Compute the shortest latency path P from node S to node D
 - 7) $P = (v_1, v_2, \dots, v_h)$, **set** $w(v_i, v_{i+1}) \rightarrow w^*$;
 - 8) $G(V, E, C) \rightarrow G'(V, E, C)$;
 - 9) **Dijkstra** $(s, d, G') \rightarrow P'(s, d, G')$; // Compute another path P'
 - 10) $G^* = P + P'$;
 - 11) **ILP** $\rightarrow \text{Min}((X-Y) \cdot W)$, $s.t. X, Y \in G^*$;
 - 12) **Return** X, Y ; // Solve the ILP equation and get the edge-disjoint path-pair
 - 13) **Explicit** X and Y , *first shared node* \rightarrow *node r*, *last shared node* \rightarrow *node e*;
 - 14) **Return** $V = \{r, e\}$
 - 15) **End**
-

The path calculation module informs the SR policy programming module of the path calculation results through PCEP [17]. Then, the SR policy programming module encodes the explicit path-pair and service functions into SRH.

3.2 Segment Routing Policy Programming

In the SRv6 [18] network, SR policies are encoded as SRH. When a packet arrives at a node, the node determines how to process the packet based on the semantics associated

with the active Segment ID (SID). In SRv6, there are three programming spaces. First, the path information composed of the Segment list can be programmed. Second, the optional TLV fields added in the message can be programmed. The third programmable space is contained in the Segment lists in IPv6 address form. The specific encapsulation format of SRv6 network is shown in Fig. 5.

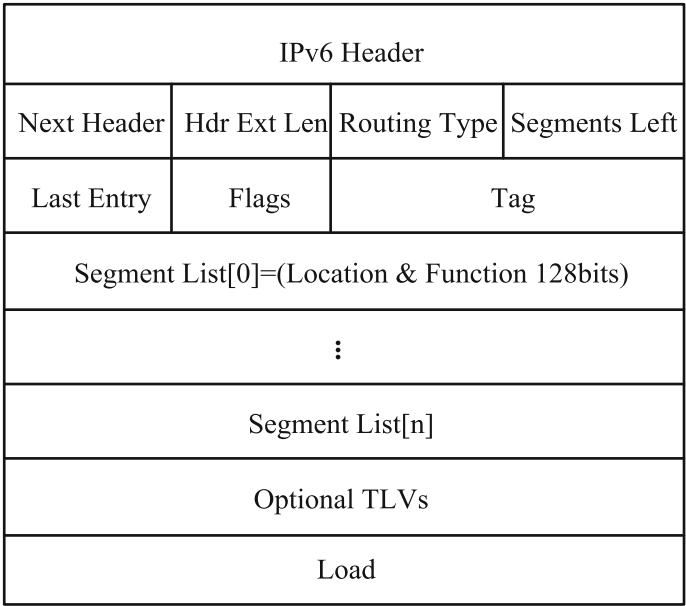


Fig. 5. Schematic diagram of SRv6 network encapsulation format

The Segment list in IPv6 address form (IPv6 SID) consists of three parts: the Locator Segment, the Function Segment and the Argument Segment. The Argument Segment is optional and can be omitted (as shown in the Segment list [0]).

After Locator part is configured, a route will be generated at the local node and diffused out through IGP [19]. Other nodes in the network can locate the node by this route. Therefore, the Locator part has routing function. Function part identifies the instruction bound to the native. If the native receives a SID, Locator part determines first, and if it is published locally, the instruction bound to the native must be processed. Instruction is a native programmable behavior, with different types of instructions. The types of SIDs are also different, different types of SIDs have different functions. The following table briefly introduces several different types of SIDs and their encoded initial registry (Table 1).

The replication and elimination functions are programmed into the Service SID. The *End. B Replication Function* pseudo-code is as follows:

Table 1. Types of SIDs and their encoded initial registry

Type	Value	Hex value	Means
End. SID	1	0 × 0001	Destination node
End. X SID	5	0 × 0005	Adjacency
End. DT SID	12	0 × 0012	Decapsulation

Algorithm 2:End.B Replication Function

Start

- 1) **If** ($NH=SRH$ & $SL>0$) **Then**
- 2) **Extract** DetNet TLV values ← SRH;
- 3) **Create** IPv6-SRH-1, IPv6-SRH-2;
- 4) **Insert** SR policy1 & TLV value1 → SRH-1, SR policy2 & TLV value2 → SRH-2;
- 5) **Delete** outer IPv6 SRH header;
- 6) **Create** Packet2 = Packet1;
- 7) **Encapsulate** packet1 → IPv6-SRH-1;
- 8) **Encapsulate** packet2 → IPv6-SRH-2;
- 9) **Set** IPv6 SA → the local address of this node;
- 10) **Set** IPv6 DA of IPv6-SRH-1 → the first segment of SR policy1;
- 11) **Set** IPv6 DA of IPv6-SRH-2 → the first segment of SR policy2;
- 12) **ELSE**
- 13) **Drop** the packet

End

3.3 Data Plane Forwarding

The data plane forwards packets based on the delivered SR policy. The pseudo-code of SR-PREOF data plane forwarding process is as follows:

Algorithm 3:SR-PREOF data plane forwarding

Start

- 1) $PCM \leftarrow$ DetNet QoS needs (Delay, jitter, bandwidth); // The path calculation module obtains deterministic service requirements
- 2) $PCM \leftarrow G(V, E)$; // PCM gets network topology parameters
- 3) $LP-EDJPP(s, G, d) \rightarrow P = \{X, Y\}$, $V = \{r, e\}$; // Path calculation algorithm
- 4) $PCM \rightarrow SRPCM$, Create SR policy; // SRPCM creates SRH
- 5) $DPFM \leftarrow$ SR policy, forward DetNet Flow;
- 6) **Execute** End.B. Replication Function in node c
- 7) **Execute** End.B. Elimination Function in node e ;
- 8) **Forward** the DetNet Flow to the DA;

End

3.4 An Example of SR-PREOF Mechanism

The SR-PREOF mechanism can be simulated by Cisco IOS XR [20]. SRv6 data plane configuration is shown in Fig. 6. The output of path calculation module is replication node R2, elimination node R9, and the explicit path-pair are P1: R2-R3-R5-R7-R9 and P2: R2-R4-R6-R8-R9.

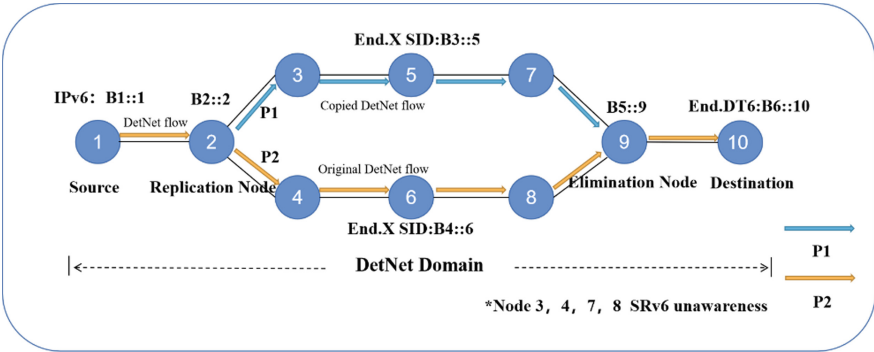


Fig. 6. Schematic diagram of the SRv6 data plane configuration

Then the operation process of the SR-PREOF mechanism is as follows:

SR Policy Encoding at the Source Node. The SR policy programming module encodes the output results of the path calculation as SRH and sends them down to the source node R1, as shown in Fig. 7.

Replication Function at Node R2. The source node sends the DetNet flow to the replication node R2 according to the Segment lists. Node R2 performs a local instruction action based on the outer IPv6 address B2::2 and then performs *End. B Replication Function*. Get two new SRHs: SRH-1 and SRH-2 as shown in Fig. 8.

Forwarding Packets in Parallel. Replicated data packets and original data packets are forwarded explicitly according to the Segment lists in SRH-1 and SRH-2. Note that R3, R4, R7 and R8 nodes without SRv6 function do not handle SRH extension headers, but they lookup IPv6 routing tables for packet forwarding.

Elimination Function at Node R9. Data packets are eliminated when the elimination node R9 hits the *End. B Elimination Function*, which is similar to the replication instruction.

Forwarding Packets to Destination Node R10. After the elimination operation at R9, the new SRH-3 is obtained. The local SID table is searched according to the outer IPv6 address B6::10. Then the *End. X SID* instruction is executed and the traffic is forwarded to the destination node R10.

Version	Traffic Class	Flow Label	
Payload Length		Next Header=43	Hop Limit
Source Address(B1::1)			
Destination Address(B2::2)			
Next Header	Hdr Ext Len	Routing Type=4	Segments Left=2
Last Entry	Flags	Tag	
Segment List[0]=(B6::10)			
Segment List[1]=(B5::9)			
Segment List[2]=(B2::2)			
Service SID=(End.B.Replication:B2::2)			
Service SID=(End.B.Elimination:B5::9)			
Optional TLV objects			
IPv6 Payload			

Fig. 7. Schematic diagram of SRv6 packet format of source node R1

Version	Traffic Class	Flow Label	
Payload Length		Next Header=43	Hop Limit
Source Address(B2::2)			
Destination Address(B3::5)			
Next Header	Hdr Ext Len	Routing Type=4	Segments Left=2
Last Entry	Flags	Tag	
Segment List[0]=(B6::10)			
Segment List[1]=(B5::9)			
Segment List[2]=(B3::5)			
Service SID=(End.B.Replication:B2::2)			
Optional TLV objects			
IPv6 Payload			

Version	Traffic Class	Flow Label	
Payload Length		Next Header=43	Hop Limit
Source Address(B2::2)			
Destination Address(B4::6)			
Next Header	Hdr Ext Len	Routing Type=4	Segments Left=2
Last Entry	Flags	Tag	
Segment List[0]=(B6::10)			
Segment List[1]=(B5::9)			
Segment List[2]=(B4::6)			
Service SID=(End.B.Replication:B2::2)			
Optional TLV objects			
IPv6 Payload			

Fig. 8. Schematic diagram of SRH-1 and SRH-2

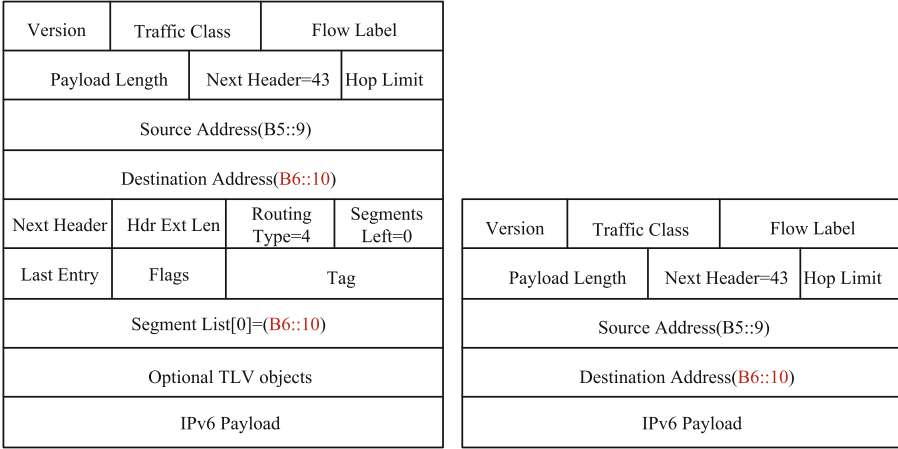


Fig. 9. Schematic diagram of SRv6 packet format of node R9 and R10

Since $SL = zero$, *End. DT6* instruction is executed at node R10, and the SRH extension header is popped to obtain the common IPv6 message, as shown in Fig. 9. The entire DetNet service protection mechanism is finished.

It can be seen from the above example that the service protection in DetNet can be implemented through the SR technology.

4 Performance Evaluation

This section presents a quantitative comparison and analysis of the proposed SR-PREOF mechanism compared with the traditional Shortest Path First (SPF) and PREOF mechanisms. The performance evaluation metrics include the path reliability, packet reception rate and E2E jitter and worst-case latency bound.

4.1 Simulation Settings

The network topology in the simulation is shown in Fig. 10. The edge-disjoint path-pair X-Y is generated according to the designed LP-EDJPP routing algorithm. $X = 1-2-3-4-5-6-7$, $Y = 1-2-8-4-10-6-7$. The replication node is node 2 and the elimination node is node 6.

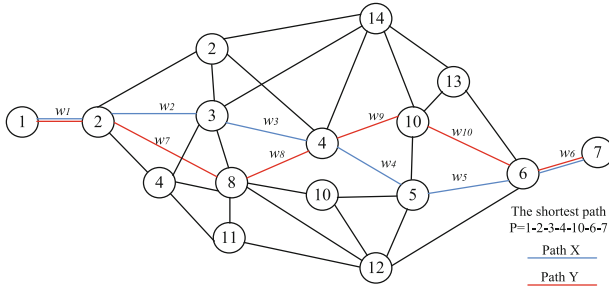


Fig. 10. Schematic diagram of the network topology used in simulation

Table 2 shows the flow model used in the simulation.

Table 2. Flow model and simulation parameters

Parameter	Distribution	Value
Link bandwidth (Gbit/s)	-	1
Probability of edge failure	Uniform distribution	(0, 0.001)
Number of packets	-	5×10^4
Size of packets (bytes)	Uniform distribution	64–500
Arrival process of packets	Poisson distribution	500
Interval of packets (microseconds)	Negative exponential distribution	10

4.2 Performance Evaluation Metrics

Path Reliability. It refers to the probability of successful packet forwarding without edge failures. The path reliability is also associated with the offered load. Equation (3) indicates the probability that each edges in the selected path 1 do not fail. Equation (4) indicates the probability that path 2 successfully forwards the packet. Equation (5) represents the path reliability of the proposed mechanism.

$$R_1 = \prod_{i,j \in V_1} (1 - pij \times offered_load) \tag{3}$$

$$R_2 = \prod_{i,j \in V_2} (1 - pij \times offered_load) \tag{4}$$

$$R_t = 1 - (1 - R_1) \times (1 - R_2) \tag{5}$$

where pij is the probability of edge failure. The offered load is between (0, 1), which represents the ratio of the current network throughput to the total network bandwidth.

Packet Reception Rate. It refers to the probability that packets can be successfully forwarded to the destination node while meeting the QoS requirements of the service (bounded latency is the requirement in the simulation). In this simulation, the network topology is generated in random with different number of nodes under the same offered load 0.7. Packet reception rate is the ratio of the number of packets arrived within the latency bound over the total number of transmitted packets. We will make a comparison between SR-PREOF, traditional PREOF and traditional Link Pruning (LP) scheme.

End-to-End Jitter and Worst-Case Latency Bound. E2E jitter refers to the variation of packet delay. The deterministic service QoS requirements differ from the traditional requirements in that the deterministic traffic forwarding focuses on the worst-case latency bound, rather than the average latency. We divide the latency into three parts here: 1) Link latency. It is given by the network topology. Latency on each link is constant. 2) Processing latency. It is related to the node computing capacity and the offered load. It is assumed that it follows random distribution from Ta to Tb . 3) Queuing latency. It occurs at each relay nodes. Its latency bound is $2T$, T indicates the cycle time, which is a constant. Equation (6) shows how to calculate latency bound, where N means the number of relay nodes.

$$Latency_{bound} = Latency_{link} + N \times (Latency_{processing} + 2T) \tag{6}$$

4.3 Simulation Results and Analysis

Figure 11 shows path reliability comparisons among SR-PREOF, PREOF and SPS mechanisms under different offered loads. With the increase of the offered load, the path reliability of three schemes all decrease. The path reliability of SPS scheme is the lowest, because it does not prepare backup paths. The proposed SR-PREOF can achieve the comparable path reliability performance compared with the traditional PREOF.

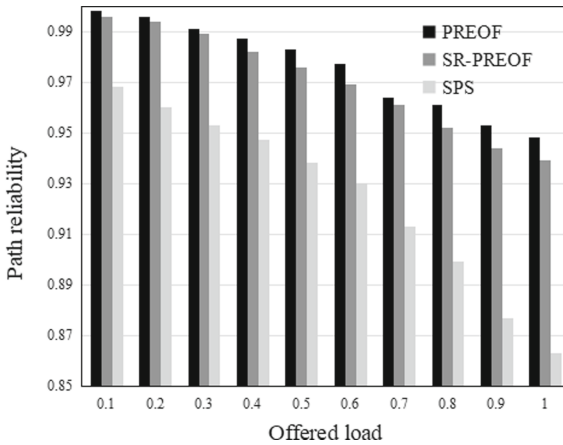


Fig. 11. Path reliability of schemes under different offered loads

As shown in Fig. 12, with the increase of the number of nodes in the network topology, the packet reception rate of the three mechanisms all increase. Moreover, the proposed SR-PREOF always has the highest packet reception rate. The advantage is obvious when the network size is small, because the other two algorithms may not find two disjoint paths. The proposed SR-PREOF increases packet reception rate by 5.6% compared with the traditional PREOF when the number of nodes is 12.

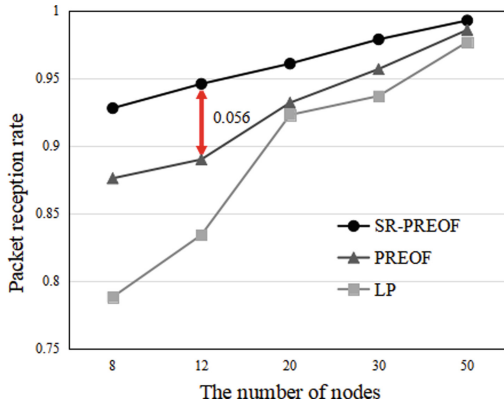


Fig. 12. Packet reception rate under different numbers of nodes

Figure 13 shows the E2E jitter and worst-case latency bound of the proposed SR-PREOF and the traditional PREOF under different offered loads. The SR-PREOF has a better jitter performance than the traditional PREOF. The latency bound of SR-PREOF is more stable than the traditional PREOF. What's more, when the offered load is 0.7, the E2E worst-case latency bound of the proposed SR-PREOF is decreased by 10.89% compared with the traditional PREOF.

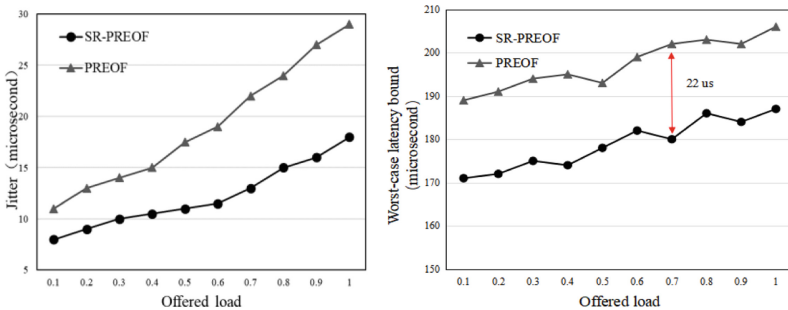


Fig. 13. End-to-end jitter and worst-case latency bound of SR-PREOF and PREOF

5 Conclusion

Based on the Segment Routing technology, this paper proposed the SR-PREOF mechanism for the service protection in DetNet. This mechanism utilizes the designed LP-EDJPP routing algorithm to obtain the equal-cost edge-disjoint path-pair for parallel packet transmissions. Simulation results show that the proposed SR-PREOF effectively improves the packet reception rate and reduces the E2E worst-case latency bound while achieving the comparable path reliability performance compared with the traditional PREOF. The packet reception rate of the SR-PREOF increases by 5.6% and the E2E worst-case latency bound decreases by 10.89% compared to the PREOF when the offered load is 0.7.

References

1. Grossman, E.: Deterministic Networking Use Cases. RFC 8578 (2019). <https://www.rfc-editor.org/rfc/pdf/rfc8578.txt.pdf>
2. Finn, N., et al.: Deterministic Networking Architecture. RFC 8655 (2019). <https://www.rfc-editor.org/rfc/pdf/rfc8655.txt.pdf>
3. Sprecher, N., Farrel, A.: MPLS Transport Profile (MPLS-TP) Survivability Framework. RFC 6372 (2011). <https://www.rfc-editor.org/info/rfc6372>
4. Koetter, R., Medard, M.: An algebraic approach to network coding. *IEEE/ACM Trans. Netw.* **11**(5), 782–795 (2001)
5. Varga, B., Farkas, J., Berger, L., et al.: Deterministic Networking Data Plane: MPLS. RFC 8964 (2021). <https://www.rfc-editor.org/rfc/rfc8964.pdf>
6. Filssils, C., Previdi, S., Ginsberg, B., et al.: Segment Routing Architecture. RFC 8402 (2018). <https://www.rfc-editor.org/pdf/rfc8402.txt.pdf>
7. Rosen, E., Viswanathan, A., Callon, R.: Multiprotocol Label Switching Architecture. RFC 3031 (2001). <https://www.rfc-editor.org/pdf/rfc3031.txt.pdf>
8. Yuchun, G., Fernando, K., Mieghem, P.: Link-disjoint paths for reliable QoS routing. *Int. J. Commun. Syst.* **16**, 779–798 (2003)
9. Huong, T.T., Khoa, N.D.D., Dung, N.X., Thanh, N.H.: A global multipath load-balanced routing algorithm based on reinforcement learning in SDN. In: 2019 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1336–1341 (2019)
10. Weiner, J., Ernst, A.T., Li, X., et al.: Solving the maximum edge disjoint path problem using a modified Lagrangian particle swarm optimisation hybrid. *Eur. J. Oper. Res.* **293**, 847–862 (2021)
11. Atallah, A.A., Hamad, G.B., Mohamed, O.A.: Routing and scheduling of time-triggered traffic in time-sensitive networks. *IEEE Trans. Ind. Inf.* **16**(7), 4525–4534 (2020)
12. Krolkowski, J., Martin, S., Medagliani, P., et al.: Joint routing and scheduling for large-scale deterministic IP networks. *Comput. Commun.* **165**, 33–42 (2021)
13. Aubry, F., Vissicchio, S., Bonaventure, O., Deville, Y.: Robustly disjoint paths with segment routing. In: CoNEXT 2018, 4–7 December 2018, Heraklion, Greece (2018)
14. Aubry, F., Lebrun, D., Deville, Y., Bonaventure, O.: Traffic duplication through segmentable disjoint paths. In: 2015 IFIP Networking Conference, pp. 1–9 (2015)
15. Qi, G.: Research on Two Disjoint Paths (In Chinese). Fudan University (2006)
16. Risald, Mirino, A.E., Suyoto: Best routes selection using Dijkstra and Floyd-Warshall algorithm. In: 2017 11th International Conference on Information & Communication Technology and System (ICTS), pp. 155–158 (2017)

17. Vasseur, J.P., Le Roux, J.L.: Path Computation Element Communication Protocol. RFC 5440 (2009). <https://www.rfc-editor.org/pdf/rfc5440.txt.pdf>
18. Filsfils, C., Camarillo, P., Leddy, J., et al.: Segment Routing over IPv6 Network Programming. RFC 8986 (2021). <https://www.rfc-editor.org/rfc/rfc8986.pdf>
19. Gross, P.: Choosing a Common IGP for the IP Internet. RFC 1371 (1992). <https://www.rfc-editor.org/pdf/rfc1371.txt.pdf>
20. Cisco policies and processes. Cisco Guide to Harden Cisco IOS XR Devices. https://tools.cisco.com/security/center/resources/increase_security_ios_xr_devices.html



Cloud-Edge Collaboration Based Power IoT Scene Perception Mechanism

Sujie Shao¹(✉), Congzhang Shao¹, Cheng Zhong², Shaoyong Guo¹,
and Pengcheng Lu²

¹ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

buptssj@bupt.edu.cn

² Xiong'an New Area Power Supply Company of State Grid Hebei Electric Power Co., Ltd., Xiong'an 071600, Hebei, China

Abstract. Fast and high-quality scene perception is an important guarantee for the efficient, stable and reliable operation of the power Internet of things, which can assist the decision-making of upper-level applications. The transmission delay of scene perception based on cloud computing is high, so it is difficult to meet the needs of real-time decision-making the mode based on edge computing is not competent for all real-time perception tasks due to the limited computing resources. For this reason, this paper proposes a scene awareness mechanism of the power Internet of things based on cloud-edge collaboration. A scene information awareness architecture based on cloud-edge collaboration is constructed, and a scene information processing flow that distinguishes dynamic instances, static instances and general instances is designed to support local scene information edge awareness and global scene cloud synthesis. Focusing on the construction of high-precision neural network recognition model of high-frequency dynamic examples, using the idea of transfer learning, a neural network model training framework based on cloud-edge collaboration is designed. Simulation results show that the scene perception mechanism proposed in this paper can effectively reduce the perception processing delay and model training time on the basis of accurately perceiving the scene, and improve the adaptability of the perception model to high dynamic scenes.

Keywords: Cloud-edge collaboration · Power IoT · Scene perception · Transfer learning · Edge intelligence

1 Introduction

In recent years, with the vigorous development of new-generation information technologies such as 5G, cloud computing, and artificial intelligence, power IoT technology

This work is supported by the Science and Technology Project of State Grid Corporation of China: Research and application of key technologies for wireless communication network coverage of urban power underground pipe gallery (Grant No. 5700-202113189A-0-0-00).

has received extensive attention and development. Scene information perception is an important cornerstone to realize the technology vision of smart power IoT technology, which involves the collection and processing of video image data and various sensor data in power IoT scenarios. In actual scenarios, all kinds of information in the Internet of Things are changing rapidly, which requires real-time and accurate perception.

Cloud computing mode is a method of power IoT scene information perception [1], which meets the needs of scene information perception to a certain extent. However, the model based on cloud computing has great limitations, which are mainly reflected in the following aspects:

- Real-time. In the actual scenario of the Internet of Things, the sensing information changed rapidly. In order to make the upper-layer applications and personnel make smooth decisions, the scene information needs to be sensed in time. However, uploading the scene information to the cloud server for processing and then returning it will result in excessive delay.
- Resource utilization. In the power IoT scenario, there is a lot of redundant and repeated information, uploading unprocessed raw data directly to the cloud computing center for processing will occupy a large amount of network bandwidth, resulting in a decrease in the efficiency of network resource utilization.
- Accuracy. The characteristics of instances in different sub-scenarios of the Internet of Things are different, to generate perception models for specific scenarios was difficult to the cloud computing mode.

As an emerging distributed computing framework [2], the edge computing model was proposed to solve the limitations of the cloud computing model, and its idea is to process data at the edge of the network. With lower processing delay and can reduce the load of the network, the edge computing mode solved the limitations of the cloud computing mode to a certain extent. However, servers deployed at the edge of the network often have limited computing resources, when the computing power required for computing tasks in the scene exceeds the server or faces some more complex computing tasks, it will not be able to meet the demands of scene information processing.

In recent years, cloud-edge collaboration has begun to receive extensive attention from researchers. It combines the advantages of cloud computing and edge computing. The main idea is to complete data preprocessing and some computing tasks on edge servers, and complete computing tasks that edge servers cannot perform on cloud servers, finally complete computing tasks quickly and with high quality.

This paper proposes a cloud-edge collaboration based power IoT scene perception mechanism. According to different types of information in the scene, a scene-aware architecture that distinguishes instance types is designed. For the construction of high-precision recognition models for high-frequency changing dynamic instances, a cloud-edge collaborative neural network model training framework is proposed, and finally of power IoT scene information.

The main contributions of this paper are as follows:

- Proposes a cloud-edge collaborative scene information perception architecture for the Internet of Things. This architecture divides a wide variety of information in the

power IoT scene into different types of instances, and then designs corresponding perception algorithms for different instance types to complete perception fusion at the edge, finally synthesize the global scene in the cloud server, reduced processing latency while ensuring the perception quality.

- Proposes a neural network model training framework based on cloud-edge collaboration is, which solves the problem of building a high-precision recognition model for high-frequency changing dynamic instances. The framework adopts the idea of migration learning, assists the training of the neural network model on the edge server with the help of the cloud server, and shares some parameters of each edge neural network model through the cloud server during the training process, which reduces the burden on the edge server and improves the convergence speed of the model in highly dynamic scenarios of the neural network model.

2 Related Work

Many literatures have carried out research on power IoT or IoT scene perception. T Kobzan et al. [10] proposed an IoT architecture scheme based on smart factory network, which transforms the current traditional network infrastructure into a more flexible SDN (Software Defined Network)-based infrastructure to provide context awareness for IoT and other services. P Hořejší et al. [11] applied ICT-enhanced solutions to smart factories and developed a new smart solution that reduced the workload of employees and improved the efficiency of factories. Wang Yaonan et al. [12] proposed a machine vision perception control system scheme based on cloud computing, which solved the contradiction between the high real-time requirements of visual control applications and the extremely high computational complexity to a certain extent. Zhang Ansi et al. [13] studied the perception method, equipment data fusion and analysis technology of digital workshop manufacturing field equipment data, which provided a guarantee for the realization of digital workshop equipment safety, stability, high efficiency and green production. The above scene perception work is based on cloud computing, which meets the needs of scene perception to a certain extent, but it is still difficult to meet the needs of actual scenes when dealing with tasks with high real-time performance.

Since edge computing was proposed, it has been widely used in various fields, especially in power IoT scenarios. W Shi et al. [2] introduced the concept of edge computing and summarized its technical challenges and application prospects. Z Zhou et al. [14] outline key techniques for deep learning models for network edge training/inference, and discuss future research opportunities for edge intelligence. Jung Woo-Kyun et al. [15] proposed an application measure of IoT smart sensors, smart IoT, and small data processing combined with edge computing, which provided a reference for SMEs to apply IoT. Z Song et al. [16] implemented an edge-based sensing system, including deep learning-based image recognition, data model retraining, etc., which effectively utilized the resources available at the edge and greatly reduced network traffic. The above-mentioned mode based on edge computing can better solve the problem of real-time task response, but limited by the computing power of edge servers, it is still difficult to perform more complex computing tasks such as deep neural network training.

In recent years, the cloud-edge collaboration model has been proposed to meet the needs of actual scenarios. This model combines the advantages of the cloud computing

model and the edge computing model, and can ensure real-time response while being competent for computing tasks. Chen Yiping [17] analyzed the realization principles and research ideas of technologies such as resource collaboration, intelligent collaboration, and application management collaboration, and pointed out the current challenges and future development directions of cloud-edge collaboration. At present, some works have applied the cloud-edge collaboration model to practical scenarios. Zhao Yu et al. [18] proposed a target detection framework combining edge computing and federated learning, which improved the detection accuracy by 18% in video surveillance scenarios and effectively reduced model training time. T Jing et al. [19] proposed a cloud-edge collaboration architecture suitable for the Internet of Things, which meets the complex network data computing and real-time analysis requirements of local data in the Internet of Things. C Ding et al. [9] proposed a cloud-edge collaboration cognitive service framework, which uses the cloud server to accelerate the training of lightweight neural networks in the edge server, improved the training speed of the model and the accuracy of cognitive services. Xu Mengwei et al. [20] proposed a new type of machine learning model training mode based on cloud-based private data of mobile terminals, which effect exceeds the traditional centralized/federated training mode, and greatly reduces the computational and energy overhead for model training. However, there is no mature research on the application of cloud-edge collaboration mode to power IoT scene perception in the existing work, and the cloud-edge collaboration mechanism needs to be further improved.

3 Cloud-edge Collaboration Based Scene Information Perception Architecture

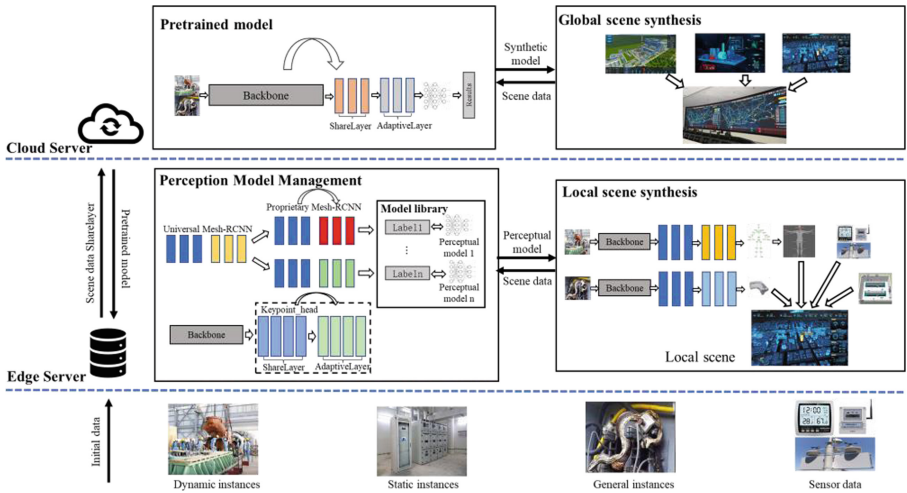


Fig. 1. Cloud-edge collaboration based scene information perception architecture

This paper designs a cloud-edge collaboration based scene information perception architecture, as shown in Fig. 1. The information involved in the power IoT scene includes sensor data (voltage, temperature, humidity, etc.), image data (personnel, inspection robots, faulty equipment, etc.). This paper divides the scene information into static instances, dynamic instances, general instances and sensor data four types. After the terminal device collects the scene information, the edge server obtains the local scene through three stages of instance extraction, instance perception and scene fusion, and then each edge server uploads the synthesized local scene to the cloud server, and finally the cloud server synthesizes the global scene, detailed process is as follows.

3.1 Instance Extraction

The instance extraction stage completes the classification of scene information. In the power IoT scene, the initial scene information includes two categories of sensor data and image data, and the local scene perception process is shown in Fig. 2.

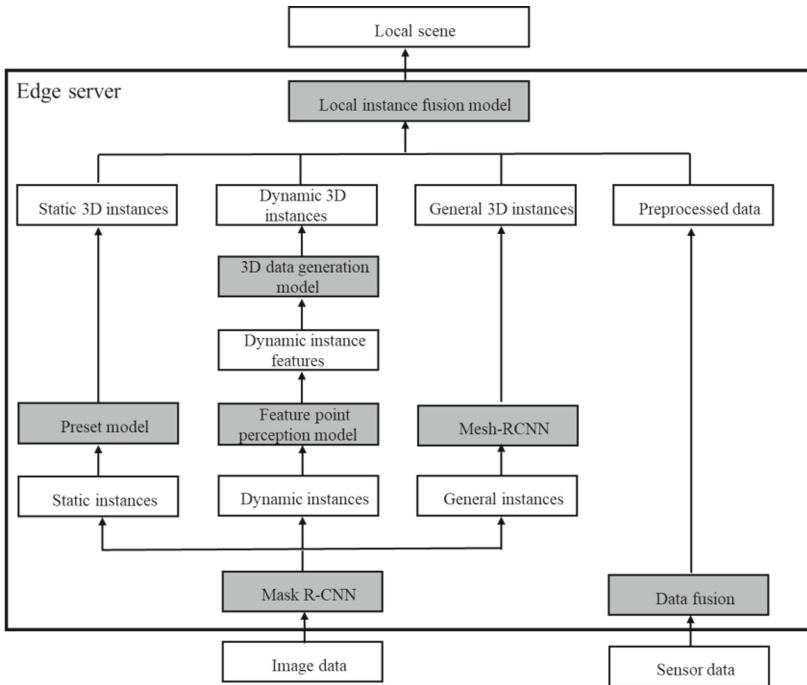


Fig. 2. Local scene perception process

In the Internet of Things scenario, there are many kinds of sensor data, which need to be fused to facilitate use. In this paper, the fusion analysis method based on small sample learning is used to fuse the sensor data [13]. Small sample learning can be divided into one-shot K-way and N-shot K-way. The difference between the two is that the number of samples in the K classes to be merged is different: N-shot K-way has N samples in each

class. Samples, while one-shot K-way has only one sample per class. Due to the variety of data formats in the power IoT scenario, it is difficult to summarize the characteristics of a certain type of data with a single sample. Therefore, this paper uses the N-shot K-way algorithm to fuse the sensor data. The specific algorithm is as follows:

For the sensing data, first construct K support set classes, each with N samples (S_1, \dots, S_N), the purpose of the sensing algorithm is to determine which support set class the collected sensor data should belong to, and the optimization goal is as follows:

$$C(\hat{x}, (S_1, \dots, S_N)) = \arg \max_i \left(\sum_{n=1}^N P(\hat{x}, x_{i_n}) \right) \quad (1)$$

where $x_{i_n} \in S_n, i \in K$. Through the N-shot K-way few-shot learning process, the sensory data in the scene are fused into K classes.

The image data is obtained by the image acquisition device, and contains a variety of instance types. In this paper, the image data is divided into three categories: static instances, dynamic instances and general instances. Mask-RCNN is one of the commonly used models for object detection, which can detect objects in images and generate masks. This paper uses the Mask-RCNN model to extract and classify the instances, and then use the corresponding perception algorithms to complete the perception processing. 3.2 introduces the detailed perception algorithms of the three instances.

3.2 Instance Perception

Static Instances. Static instances are objects with relatively fixed characteristics in power IoT scenarios, including warehouses, generator sets, and underground pipe galleries. The scene information perception architecture proposed in this paper deploys the 3D model of static instances in the edge server in advance. After the instance extraction stage Mask-RCNN identifies the static instances in the scene, it directly calls the preset 3D model to participate in the fusion of local scenes.

Dynamic Instances. Dynamic instances are objects whose characteristics change in real time in power IoT scenarios, such as workers and inspection robots. For this type of instance, this paper uses the method of first detecting key points, and then synthesizing the model based on the parameters of the key points to generate 3D models. Taking the generation of a 3D model of the human body as an example, the `keypoint_head` branch of Mask RCNN is used to detect the parameters of the key points of the human body, and then the detected parameters are input into the SMPL parameterized human body model to synthesize the 3D model of the human body to complete the perception of the human body instance. In order to further improve the perception accuracy of dynamic instances, this paper will also train the dynamic instance perception model in real time based on the collected scene information. The detailed training method will be introduced in Sect. 4.

General Instances. A general example is abnormal information in power IoT scenarios, such as illegal intrusions, faulty devices, etc., which are characterized by unpredictable characteristics. The 3D model of this type of instance can only be directly perceived from the image information. In this paper, Mesh-RCNN is used to perceive general instances.

Mesh-RCNN is improved from Mask-RCNN. It is a model after adding mesh branches on the basis of Mask-RCNN, which can directly generate 3D models from two-dimensional image data. After Mask-RCNN identifies general instances in the initial image, the sub-images are segmented through masks, and then Mesh-RCNN is used to generate mesh 3D data of instances based on the sub-images. In actual scenarios, general instances with similar characteristics may appear multiple times in a period of time. In order to improve the perception accuracy of general instances, this paper trains the corresponding Mesh-RCNN models for different types of instances.

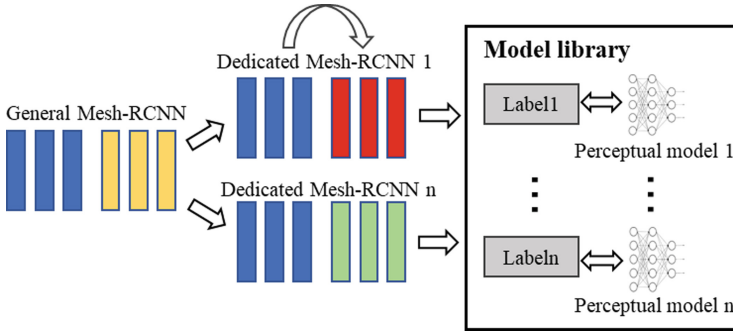


Fig. 3. Mesh-RCNN training differentiation process

Figure 3 shows the training differentiation process of Mesh-RCNN. This process is mainly completed on the edge server. First, the Mask-RCNN is trained based on the labeled data, so that it can detect several specific types of general instances, and label labels for different types, and then train the corresponding Mesh-RCNN for each label.

Under the initial conditions, the Mesh-RCNN model corresponding to each label is initialized as a general Mesh-RCNN model pre-trained on the cloud server, and then the cloud-edge collaborative perception architecture begins to perceive general instances in the scene information and collect general instance data in real time. When the instance data corresponding to a certain label is accumulated to a certain amount, the mesh label of the instance data will be obtained through manual labeling, and the Mesh-RCNN model corresponding to the label will be trained. When the general instance corresponding to a label is detected again, the Mesh-RCNN model that has been specifically trained will be called.

3.3 Scene Fusion

After the instance perception phase is completed, the 3D models of static instances, dynamic instances, and general instances will be fused at the edge server, and then the sensor data will be matched to the 3D scene, therefore obtains the local scene. The merged local scene will retain two copies: one is saved on the edge server to provide real-time services to users, and the other is uploaded to the cloud server, where the global scene is synthesized.

Denote the scene initial data as Ori, sensor data as Sen, dynamic instance as Dym, static instance as Sta, general instance as Gen, local scene perception result as Loc, global scene perception result as Gol; instance extraction process as F, instance perception The process is P and the scene fusion process is B. From the perspective of data flow, the perception process of local scenes can be expressed as the following formula:

$$Ori \xrightarrow{F} Sen, Dym, Sta, Gen \xrightarrow{(P_{sm}+P_{Dmm}+P_{com}+P_{sen})+B} Loc \quad (2)$$

The synthesis process of the global scene can be expressed as the following formula:

$$Loc_1, Loc_1, \dots, Loc_n \xrightarrow{C} Gol \quad (3)$$

In the above-mentioned scene information perception architecture based on cloud-edge collaboration, after the perception model is obtained, the scene information will be directly perceived by the edge server, and only the generation and training process of the perception model involves the collaborative work of the cloud server and the edge server. In the generation and training of the instance-aware model, the sensing data and static data-aware models are fixed models that are preset on the edge server, although the general instance needs to download the perception model from the cloud server, the training process on the edge server is independent, none of the three involve the collaboration of cloud servers and edge servers. In fact, only the training of dynamic instance-aware models involves cloud-edge collaboration, and the part where the scene perception speed and quality can be optimized is also here. Therefore, this paper will focus on the optimization of dynamic instance-aware model training methods, and the details will be introduced in Sect. 4.

4 Neural Network Model Training Framework Based on Cloud-edge Collaboration

A typical task of dynamic instance perception in power IoT scenarios is the detection of human pose key points, this paper takes this as an example to study the dynamic instance perception algorithm. To this end, a neural network model training framework based on cloud-edge collaboration is proposed, as is shown in Fig. 4. The core idea is to train the neural network perception model for specific scenarios on the edge server, and pre-train the neural network model with the help of the cloud server to reduce the load of edge server.

Figure 4 shows the cloud-edge collaborative training framework proposed in this paper. A deep neural network model, called CloudNet, is deployed on the cloud server; a lightweight neural network model, called EdgeNet, is deployed on the edge server, which is obtained by reducing some of the backbone network layers by CloudNet. The EdgeNet backbone network parameters are the same as those of the corresponding remainder in CloudNet. In real-time training, in order to obtain a human pose perception model for a certain scene, this paper trains EdgeNet based on the scene information collected in real time.

The paper [22] pointed out that the higher layers of the neural network can extract the general features of the data. In the human pose keypoint recognition task, the features of

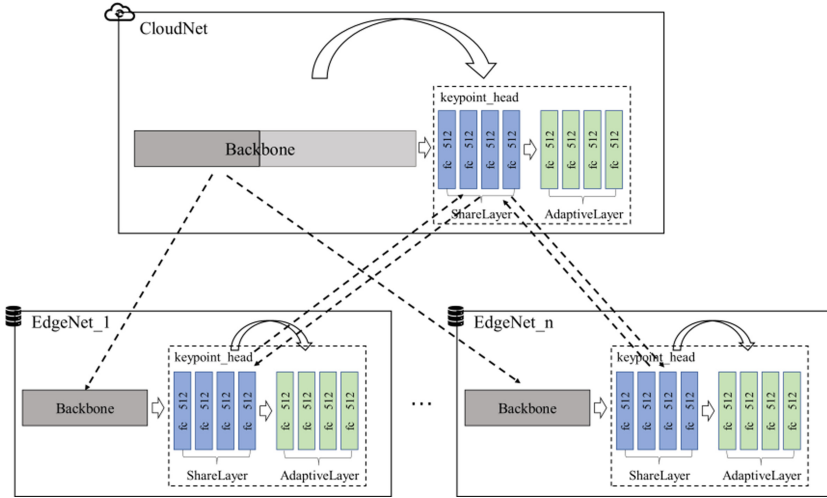


Fig. 4. Neural network model training framework based on cloud-edge collaboration

human pose overlap in each sub-scenario. For this reason, this paper sets the high m layer of the keypoint_head part of the ROI head of EdgeNet as ShareLayer, which is used to extract the common features of the instances in each sub-scenario, and its parameters are shared by each EdgeNet; meanwhile set the rest of the keypoint_head as AdaptiveLayer, which is different in each EdgeNet to adapt to the different characteristics of instances in each sub-scenario.

Regarding the acquisition of labels required for EdgeNet training, this paper uses CloudNet to automatically generate and obtain labels required for training. This is because CloudNet has a large number of neural network layers and can obtain higher recognition accuracy. CloudNet is distributed to the edge server during the initialization phase, which is recorded as Label-builder. Compared with the traditional manual labeling method, automatic label generation with CloudNet can reduce the time required for label creation and speed up model training and updating.

The training process of EdgeNet can be divided into three stages: EdgeNet initialization, EdgeNet learning, and Sharelayer update. The detailed process is as follows:

4.1 EdgeNet Initialization

In the EdgeNet initialization phase, a large amount of data is used to pre-train CloudNet on the cloud server, and then some of its parameters are sent to the edge server to assist EdgeNet training.

Write W_c as the parameter of CloudNet, W_{c-b} is the parameter of the part shared with EdgeNet in the CloudNet backbone network, W_{c-r_s} is the parameter of ShareLayer in the ROI head of CloudNet; W_e is the parameter of EdgeNet, W_{e-b} is the parameter of the EdgeNet backbone network, W_{e-r} is the EdgeNet ROI head. The parameters of W_{e-r_s} and W_{e-r_a} are the parameters of ShareLayer and AdaptiveLayer in the ROI head

of EdgeNet respectively, and N is the data volume of the training set. Given a training set $\{x_i, y_i\}_{i=1}^N$, EdgeNet will be trained to optimize the following loss function:

$$f(W_e) = \frac{1}{N} \sum_{i=1}^N \theta(y_i, \sigma(f(x_i; W_e))) \quad (4)$$

After the edge server receives the CloudNet sent by the cloud server, it first separates W_{c-b} and W_{c-rs} in CloudNet as the parameters of W_{e-b} and W_{e-rs} in EdgeNet, then initializes W_{e-ra} to a random value, and then combines the three parts of, W_{e-b} , W_{e-rs} and W_{e-ra} to form W_e , that is, the parameters of EdgeNet; then optimize W_e by formula (4) to complete the training of EdgeNet. This paper uses CloudNet to assist EdgeNet training, freezes the parameters of the backbone network in EdgeNet during the training process, and only fine-tunes the parameters of the ROI head. Use W_e^* to represent the updated value of W_e , and the connector \cup to connect the two parameter sets. The detailed steps of EdgeNet initialization are shown in Algorithm 1.

input : $\{x_i, y_i\}_{i=1}^N, W_c, \text{CloudNet}$

output: EdgeNet, Label-builder

1. The cloud server sends CloudNet (parameter value W_c) to the edge server;
2. The edge server processes CloudNet to obtain three copies of W_c, W_{c-b}, W_{c-rs} ;
3. Edge servers build EdgeNet;
4. Initialize $W_{e-b}, W_{e-rs}, W_{e-ra}$ of EdgeNet to W_{c-b}, W_{c-rs} , random values respectively;
5. $W_e = W_{e-b} \cup W_{e-rs} \cup W_{e-ra}$;
6. $W_e^* \leftarrow \arg \min_{W_e} f_e(W_e)$;
7. Return EdgeNet as W_e^* and Label-builder as W_c

Algorithm 1. EdgeNet initialization

4.2 EdgeNet Learning

In the EdgeNet initialization phase, EdgeNet completed the initial training with the assistance of CloudNet. In actual scenarios, the edge server will continuously receive the instance information collected by the terminal from the scene. The paper [21] pointed out that real-time information is very important to improve the quality of the model. To this end, this paper further trains EdgeNet using real-time collected instance information. However, the initially collected instance information has no labels. To solve this problem, this paper draws on the method in [9] and uses the Label-builder obtained in the EdgeNet initialization phase to generate the labels required for training. After the edge server receives the instance information, it first uses the EdgeNet to obtain the perception results. Whenever the number of newly collected instances accumulates to M , the Label-builder

will be used to generate the labels of the instances to retrain the EdgeNet.

$$f(W_e^*) = \frac{1}{M} \sum_{i=1}^M \theta(y_i, \sigma(f(x_i; W_e^*))) \quad (5)$$

Since Label-builder can automatically generate labels in real time, EdgeNet learning is continuous, and its detailed process is shown in Algorithm 2.

- input** : EdgeNet, W_e^* , instance $\{x_i\}_{i=1}^M$, Label-builder
output: Updated EdgeNet
1. Edge server uses Label-builder to get label $\{y_i\}_{i=1}^M$ of instance $\{x_i\}_{i=1}^M$;
 2. $W_e^{**} \leftarrow \arg \min_{W_e} f_e(W_e)$, W_e^{**} is the updated value of W_e^* ;
 3. Returns the value of EdgeNet as W_e^{**}

Algorithm 2. EdgeNet learning

4.3 Sharelayer Update

As EdgeNet learning continues, EdgeNet is constantly adapting to changing scene information. To further accelerate the convergence speed of EdgeNet learning and extract common features of instances in different scenes, this paper sets a shared layer ShareLayer(W_{e-rs}) in the ROI head of EdgeNet (as described in the beginning of Sect. 4). In real-time training, each edge service will extract the parameters of Sharelayer in EdgeNet and upload it to the cloud server after retraining the EdgeNet on it for a certain number of times. After collecting all the parameters of the Sharelayer in EdgeNet, the cloud server will use the Fedavg [23] algorithm to summarize the gradients generated by the Sharelayer in each EdgeNet during training, update the Sharelayer with the aggregated gradient and send it to each edge server.

Write W_{c-rs}^* as the updated parameter of Sharelayer. The detailed update process of Sharelayer is shown in Algorithm 3:

input : N EdgeNets, training batch size n_i , W_{e-rs_i} , W_e , $i = 1, 2, \dots, N$, Sharelayer copy on cloud server, W_{c-rs}

output: Updated ShareLayer

1. Each edge server separates the Sharelayer parameter (ie. W_{e-rs_i}) from EdgeNet and uploads it to the cloud server;
2. Calculate the gradient generated by each EdgeNet
 $g_i = W_{e-rs_i} - W_{c-rs}$, $i = 1, 2, \dots, N$;
3. Calculate the weighted average gradient sum $g = \sum_{i=1}^N \frac{n_i}{n} g_i$, where $n = \sum_{i=1}^N n_i$;
4. $W_{c-rs}^* \leftarrow W_{c-rs} - \eta(g + \partial(W_{c-s}))$;
5. Returns the updated ShareLayer as W_{c-rs}^* , and sends it to each edge server;
6. Edge servers update their ShareLayer in EdgeNet to W_{c-rs}^*

Algorithm3. Sharelayer update

After the edge server sets the ShareLayer in the respective EdgeNet to W_{c-rs}^* , it will call Algorithm 2 again to retrain the EdgeNet, which realizes further optimization of the EdgeNet. Sharelayer update and EdgeNet learning alternate in real time, but Sharelayer update will be started after a certain round of EdgeNet learning, until Sharelayer generates a more obvious gradient, so the calling frequency is slightly lower (Fig. 5).

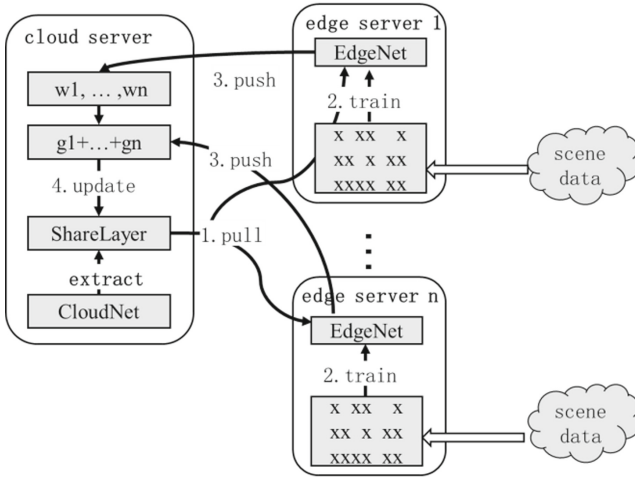


Fig. 5. Synchronous blocking parameter summary scheme

In general, the cloud-edge collaborative training framework proposed in this paper has two advantages: First, the framework can provide fast response services. This is because the edge server is closer to the user, and the number of EdgeNet network layers used for perceptual computing is less, and the computing time is short; secondly, the framework can better adapt to changes in dynamic instance characteristics of the scene.

This is because the framework trains EdgeNet based on real-time scene instance information, and introduces Label-Builder and ShareLayer so that training updates can be continuously performed in real-time and adapt to changes in scene information more quickly.

5 Simulation

5.1 Settings

Human posture key point detection is a key and representative task in the scene perception of power IoT. This paper takes this task as an example to verify the effect of the proposed cloud-edge collaborative power IoT scene perception mechanism.

The CPU model of the server used in the simulation experiment is Xeon E5-2600 v2@2.1 GHz, and the running memory size is 62.5 GB. The human pose keypoint detection framework improves from the Dectectron2 framework (based on the Mask RCNN model) commonly used in object detection. CloudNet and EdgeNet are improved from the `keypoint_rcnn_R_50_FPN_3x` model provided by Dectectron2. The detailed network structure is shown in Fig. 6 Among them, the original CloudNet model has the same structure. The backbone network of EdgeNet removes the `res4` and `res5` parts of the model, and the rest is the same as the original model. ShareLayer is the upper 4 layers of Mask RCNN's FPN, Box Head, Box_predictor and `keypoint_head`, and AdaptiveLayer is the lower 4 layers of `keypoint_head`. Simulate real-time scene data with images from COCO2017 annotated with human pose keypoints.

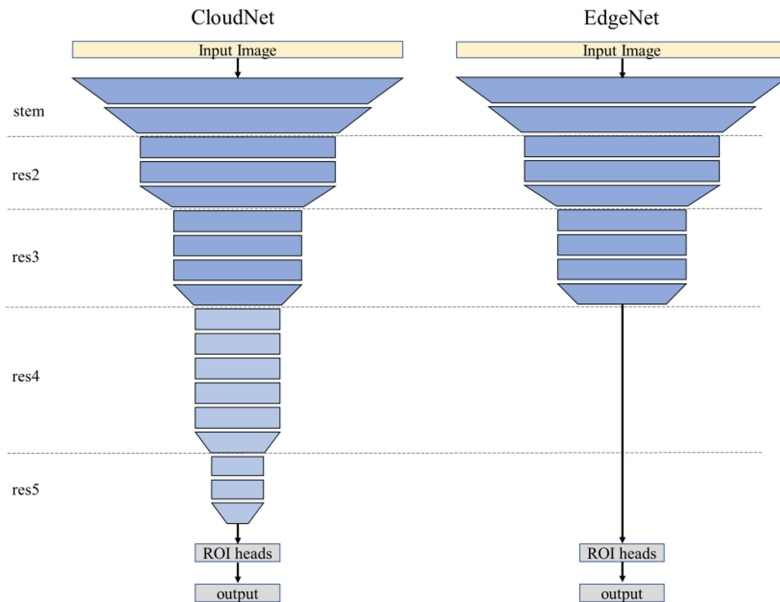


Fig. 6. Structure of EdgeNet and CloudNet (based on `keypoint_rcnn_R_50_FPN_3x`)

5.2 Results

Figure 7 shows the comparison of perception and transmission delays for the human pose keypoint detection task using the two mechanisms. Compared with the cloud computing model, the cloud-edge collaboration mechanism proposed in this paper reduces the perception delay by 39.8%. This is because EdgeNet is obtained by removing the backbone network res4 and res5 from the keypoint_rcnn_R_50_FPN_3x model, and the number of network layers is less; the transmission delay is reduced increased by 26.9% because edge servers are closer to users. The total processing delay obtained by the combined calculation of the two is reduced by 38.6%, which indicates that the collaborative perception architecture proposed in this paper can effectively improve the perception speed of scene information.

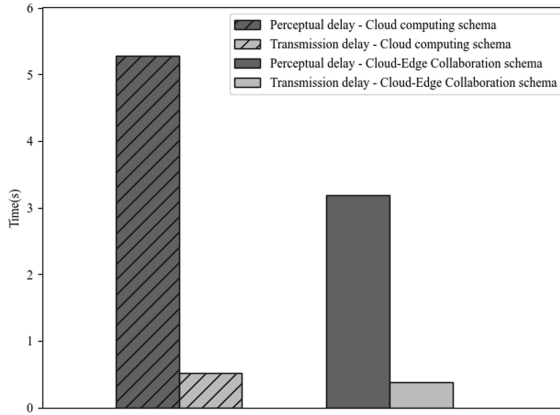


Fig. 7. Comparison of perception delay and transmission delay between the two schemas

In order to investigate the improvement of the model training effect of the proposed cloud-edge collaborative training framework, this paper simulates the process of collecting data and training EdgeNet in the actual scene. The specific settings are as follows: Under the initial conditions, the pre-trained CloudNet is deployed in the cloud server, and the training starts. At first, Algorithm 1 is called to initialize EdgeNet based on the preset 100 image data. Then simulate the process of real-time training EdgeNet in the actual scene, take 100 images as a batch, load 12 batches of pictures, and call algorithm 2 to retrain EdgeNet after each batch is loaded, and each edge server is loaded separately. After three batches of images, Algorithm 3 will be called to summarize the gradients generated by ShareLayer in the first two EdgeNet training sessions and retrain EdgeNet based on the updated gradients. That is, in the loading of every three batches of images, the first two calls to Algorithm 2 retrain EdgeNet, and the last call to Algorithm 3 to retrain EdgeNet.

Loss_keypoint is the error between the output value of Mask-RCNN on the training set and the real value. This paper uses this as a reference to measure the adaptability of the model to the new information features in the scene. Figure 8 compares the change of loss_keypoint with the number of iterations in the training process when the loading

batches are 2, 3, 5, 6, 11, and 12, and the ShareLayer in EdgeNet is set to be the upper 0, 4, and 8 layers of keypoint_head, respectively. Among them, ShareLayer is the first 4 layers of keypoint_head, which is the training method proposed in this paper. ShareLayer is the first 0 layers of keypoint_head, that is, the classic transfer learning method without ShareLayer. ShareLayer is the first 8 layers of keypoint_head, that is, the parameters of EdgeNet on each edge server are completely shared. When the loading

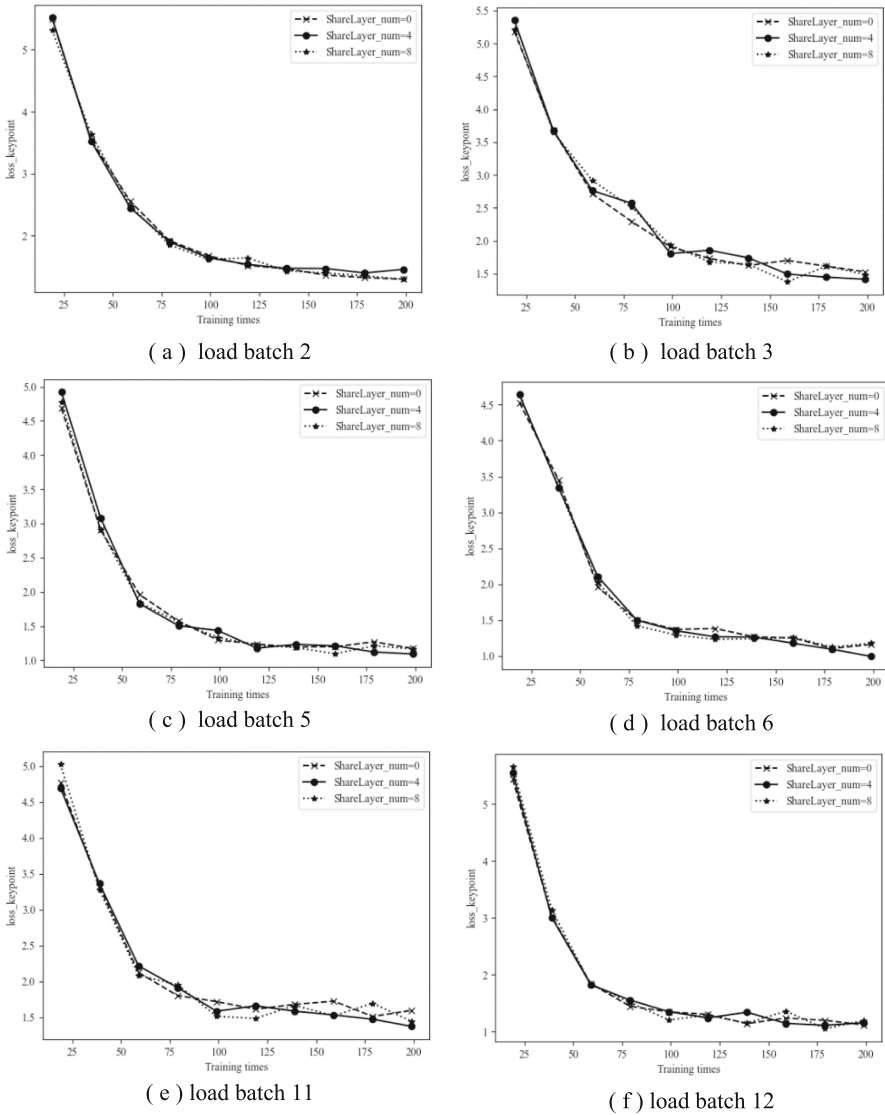


Fig. 8. The change of loss_keypoint with the number of training when the loading batches are 2, 3, 5, 6, 11, 12 and shareLayer is set as the first 0, 4, and 8 layers of keypoint_head.

batch is 2, the EdgeNet of the three structures starts to converge when the number of iterative training is about 130. At this time, the algorithm 3 synchronization ShareLayer is not enabled, and the final convergence value of loss_keypoint is also similar; when the loading batch is 3, it is enabled for the first time. Algorithm 3 synchronizes the Sharelayers of each EdgeNet. After 200 iterations of training, setting ShareLayer_num = 4 converges to the smallest loss_keypoint value, which is 4.856% lower than setting ShareLayer_num = 8, and 7.428% lower than setting ShareLayer_num = 0. When loading batches of 5, 6, 11, and 12, setting ShareLayer_num = 4 also obtained the smallest loss_keypoint. This is because, on the one hand, the setting of ShareLayer can extract the common features of each sub-scenario [22] and is jointly trained by each EdgeNet, so the convergence speed is faster; on the other hand, the 4-layer AdaptiveLayer trained by each EdgeNet allows EdgeNet to learn different the unique characteristics of sub-scene enhance the adaptability of the model to different scenes, so it achieves better results than setting ShareLayer to 8. It is worth noting that the data set of the simulated collection data and training process in this paper is randomly divided from COCO2017, and the distribution of information features of each sub-scenario is relatively uniform. If the characteristics of the datasets used in each EdgeNet training are quite different, the collaborative training algorithm proposed in this paper will have a more prominent performance.

Figure 9 shows the comparison of the training time of ClouNet, EdgeNet0, and EdgeNet. EdgeNet0 has the same network structure as EdgeNet, but the parameters are randomly initialized during training, and all parameters are adjusted in the process. As can be seen from the figure, the training time of EdgeNet is reduced by 37.9% compared to CloudNet, because the number of layers of the EdgeNet backbone network is small; the training time of EdgeNet0 is reduced by 23.2%, this is because EdgeNet only adjusts the parameters of the ROI head part during the training process, which reduces the amount of computation. This shows that the neural network model training framework based on cloud-edge collaboration proposed in this paper can effectively reduce the training time of the model.

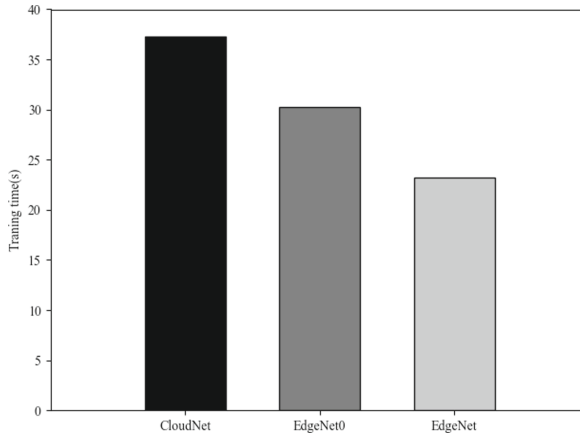


Fig. 9. ClouNet, EdgeNet0, EdgeNet training time comparison

6 Conclusion

In this paper, a cloud-edge collaboration perception mechanism is proposed for the power IoT scene perception problem. The various types of data in the power IoT scene are summarized into four categories: sensor data, static instances, dynamic instances, and general instances, and corresponding algorithms are designed to complete the perception. Dynamic instance perception is a key link in scene perception. This paper takes the task of human pose key point detection as an example to study and design a cloud-edge collaborative neural network model training framework. EdgeNet is trained based on real-time scene information to adapt to the dynamic changes of instance features in the scene, and during the training process, CloudNet is assisted and the ShareLayer shared by each EdgeNet is introduced to improve the model training effect. The simulation results of the prototype system show that the cloud-edge collaborative perception mechanism proposed in this paper can realize the rapid perception of power IoT scene information while ensuring the perception quality. The proposed cloud-edge collaborative neural network model training framework can effectively speed up the convergence of model training and adaptability to new scenarios.

Limited by the experimental conditions, this paper does not verify the proposed cloud-edge collaborative perception mechanism in actual scenarios, and the perception algorithms for other types of instances other than dynamic instances also need to be verified. In addition, there is also a lack of a dynamic allocation mechanism for cloud-side server computing resources. Future work will further improve the cloud-edge collaboration mechanism, and test and verify the effect of the mechanism in actual scenarios.

References

1. Chen, Y.: Research on 3D scene perception technology of mobile robot based on convolutional neural networks [Master's Thesis]. Harbin Institute of Technology, Harbin, 2018(in Chinese with English abstract)
2. Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L.: Edge computing: vision and challenges. *IEEE Internet Things J.* **3**(5), 637–646 (2016). <https://doi.org/10.1109/JIOT.2016.2579198>
3. Shaoguang, Y., Bingyi, F., Zengsen, W., Guo Chengyu, Q., Xiuchao.: China research on key technologies and applications of cloud-edge collaboration. *Inf. Commun. Technol.* **14**(04), 31–36 (2020). (in Chinese with English abstract)
4. Wu, G., et al. MECCAS: collaborative storage algorithm based on alternating direction method of multipliers on mobile edge cloud. In: 2017 IEEE International Conference on Edge Computing (EDGE) Honolulu, pp. 40–46 (2017)
5. Li, Y., Xu, L.: The service computational resource management strategy based on edge-cloud collaboration. In: 10th International Conference on Software Engineering and Service Science (ICSESS) Beijing, pp. 400–404 (2019)
6. Meng, C.: Research on Content Placement and Cache-aware Scheduling for Streaming Media in Cooperative Cloud-Edge Environment[Master's Thesis]. Wuhan University of Technology, Wuhan (2019). (in Chinese with English abstract)
7. Zhou, C., Lin, Z., Du, C., Wang, Z., Li, F.: Research on key technologies of tunnel robot based on cloud edge collaboratio. In: 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT) Jeju, pp. 661–666 (2020)

8. Zhang, Y., Wang, X., He, J., Xu, Y., Zhang, F., Luo, Y.: A transfer learning-based high impedance fault detection method under a cloud-edge collaboration framework. *IEEE Access*. **8**, 165099–165110 (2020)
9. Ding, C., Zhou, A., Liu, Y., Chang, R., Hsu, C., Wang, S.: A cloud-edge collaboration framework for cognitive service. *IEEE Transactions on Cloud Computing* **10**, 1489–1499 (2020)
10. Kobzan, T., Heymann, S., Schriegel, S., Jasperneite, J.: Utilizing SDN infrastructure to provide smart services from the factory to the cloud. In: *Proceedings of the 15th IEEE International Workshop on Factory Communication Systems (WFCS)*, Sundsvall, pp. 1–4 (2019)
11. Hořejší, P., Novikov, K., Šimon, M.: A smart factory in a smart city: virtual and augmented reality in a smart assembly line. *IEEE Access*. **8**, 94330–94340 (2020)
12. Yaonan, W.A.N.G., Tiejian, C.H.E.N.: Key technologies of machine vision perception and control for smart factory. *ZTE TECHNOL. J.* **22**(05), 26–30 (2016). (in Chinese with English abstract)
13. Zhang, A.: *Research on data perception, fusion and visualization in digital workshop*[PhD thesis]. Guizhou University, Guiyang, 2019(in Chinese with English abstract)
14. Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., Zhang, J.: Edge intelligence: paving the last mile of artificial intelligence with edge computing. *Proc. IEEE* **107**(8), 1738–1762 (2019)
15. Jung, W.-K., et al.: Appropriate smart factory for smes: concept, application and perspective. *Int. J. Precis. Eng. Manuf.* **22**(1), 201–215 (2020). <https://doi.org/10.1007/s12541-020-00445-2>
16. Song, Z., Cheng, J., Chauhan, A., Tilevich, E.: Pushing participatory sensing further to the edge. In: *2019 IEEE International Conference on Edge Computing (EDGE)*, Paris, pp. 24–26 (2019)
17. Chen, Y., Liu, W., Cheng, H.: Survey of cloud-edge collaboration. *Comput. Sci.* **48**(03), 259–268 (2021). (in Chinese with English abstract)
18. Yu, Z., Jie, Y., Miao, L., Jinlong, S., Guan, G.: Federated learning based intelligent edge computing technique for video surveillance. *J. Commun.* **41**(10), 109–115 (2020). (in Chinese with English abstract)
19. Jing, T., shen, J., Jia, T., Yutong, J., Ning, Z.: Application of cloud edge collaboration architecture in power IoT. In: *2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, Chongqing, pp. 18–22 (2020)
20. Xu, M., Liu, Y., Huang, K., Liu, X., Huang, G.: Autonomous learning system towards mobile intelligence. *J. Softw.* **31**(10), 3004–3018 (2020) (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4091.htm>
21. Sun, C., Shrivastava, A., Singh, S., Gupta, A.: Revisiting unreasonable effectiveness of data in deep learning era. In: *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, pp. 843–852 (2017)
22. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks?. In: *2014 Neural Information Processing Systems (NIPS)*, Montreal, pp. 3320–3328 (2014)
23. McMahan, H., Moore, E., Ramage, D., Hampson, S., Agüera y Arcas, B.: Communication-efficient learning of deep networks from decentralized data. In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*. Seattle, US (2017)



Task Allocation Mechanism of Power Internet of Things Based on Edge Routing Optimization

Cheng Zhong^{1(✉)}, Yi Li², Sujie Shao², Shaoyong Guo², and Qianjun Wang²

¹ Xiong'an New District Power Supply Company of State Grid Hebei Electric Power Co., Ltd., Xiong'an 071600, Hebei, China
1094642873@qq.com

² State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
buptssj@bupt.edu.cn

Abstract. With large number of smart devices accessing the power Internet of Things, cooperative edge computing has become a new computing paradigm for delay sensitive business. However it still faces some challenges today: 1) task allocation in edge networks, 2) route selection during data transmission. This paper proposes a cooperative computing mechanism based on edge routing optimization for power Internet of Things to solve these problems. Firstly, establish a task allocation model based on multi-edge-node cooperation in which the data routing process includes three stage: sending route, merging route and returning route. The model is designed to minimize the average task completion delay. Secondly, Improved Biogeography-based Optimization in Task Allocation and Sending Route Selection (IBBO-TASRS) algorithm is proposed to solve the task allocation and route selection. The simulation results show that the proposed algorithms can reduce the service delay well.

Keywords: Power Internet of Things · Cooperative edge computing · Task allocation · Route optimization

1 Introduction

With the development of the power Internet of Things [1], many new business terminals are emerging and connected to the sensor network [2]. For example, the remote monitoring business uses cameras to collect video data of power scenes and perform real-time intelligent analysis. Smart Augmented Reality (AR) glasses sense the operation and maintenance site and guide the operation. Unmanned Aerial Vehicles (UAV) conduct all-round inspections of remote

Supported by the Science and Technology Project of State Grid Corporation of China: Research and application of key technologies for wireless communication network coverage of urban power underground pipe gallery (Grant No. 5700-202113189A-0-0-00).

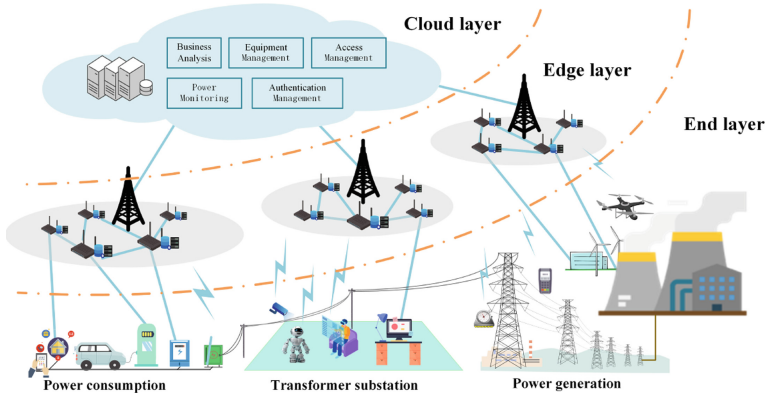


Fig. 1. The power Internet of Things architecture based on edge computing

transmission lines. The extensive deployment of a large number of sensing devices and control systems has improved the holographic perception capabilities of the power grid and the operating efficiency of power equipment. However, the reliability and safety of these service terminals have aroused widespread concern, and it is necessary to prevent attackers from eavesdropping and interception data and ensure work efficiency in complex environments. At the same time, these applications usually require a lot of computing resources and require high response delays. Wearable devices such as AR glasses usually require 10–50 ms. Today, edge computing [3], as an extension of cloud computing [4], has become an effective solution for delay and resource-sensitive services. Figure 1 shows the architecture of the power Internet of Things based on edge computing. The end layer is divided into power consumption area, power transformation area and power generation area, covering video monitoring, equipment detection, charging pile, intelligent patrol inspection, smart home and other businesses. The edge layer deploys edge nodes with computing and storage resources, such as wireless access points, routers, edge servers, etc. Clouds far away from business terminals have a large amount of computing resources, while providing business analysis, equipment management and other functions. Power business terminals are connected to edge nodes by means of wired, Wi-Fi, micro-power wireless, 4 G/5 G, and low-power wide area networks, etc. The computing tasks are uploaded to edge nodes to be completed nearby, which reduce network delay and cloud center load.

Considering the limited computing and storage resources of edge nodes, it is usually difficult for a single node to meet the demanding resource and delay requirements of service terminals. Therefore, it is necessary to adopt the cooperative edge computing [5] mode to calculate tasks initiated by service terminals. [6] has shown that cooperative edge computing is better than single node computing. However, most of the existing cooperative edge computing literature only considers the problem of task allocation without route optimization into the research [7], such as data sending route, returning route, etc. However, most of the literatures on collaborative edge computing do not consider the opti-

mization of data routing in edge networks, such as task sending routing and computation result aggregation routing, and different routing choices will have a great impact on the delay of task transmission. Therefore, in collaborative edge computing, not only the choice of computing nodes should be considered, but also the routing scheme should be paid attention to in order to reduce the delay in both transmission and computation processes. This complicates the problem of collaborative computing.

To solve the problem of task allocation and routing selection in cooperative edge computing, this paper proposes a cooperative computing mechanism based on edge routing optimization for power Internet of Things, which optimizes task allocation and route selection at the same time, and minimizes the average task completion delay while meeting business resource requirements and delay requirements. The specific contributions are as follows:

- 1) In order to reduce the delay of business completion in the power Internet of Things scenario, a task allocation model based on multi-edge-node cooperation is proposed. In the model, business request and resource models, delay models and data route models are established respectively. The data route model includes three stages: sending route, merging route and returning route. Finally, the problem of minimizing the average task completion delay under the condition of satisfying the business resource demand and delay demand is put forward.
- 2) The IBBO-TASRO algorithm is proposed to solve the task allocation and route selection. In IBBO-TASRO the elite retention strategy and improved migration probability are used to improve the biogeographic algorithm to avoid the problem of task assignment getting into local optimum.

2 Related Work

With the development of the Internet of Things in the future, the number of terminals will show explosive growth while creating more data. The cloud computing model will bring about big data transmission, privacy and energy consumption issues [8,9]. Cooperative edge computing uses multiple edge nodes to contribute resources to participate in task computing, solving the problem of insufficient resources of a single node and high latency of cloud computing. The current research on cooperative edge computing includes task allocation, route selection and other issues.

The task allocation mainly studies how tasks are distributed among the edge nodes and cloud centers to achieve optimal latency or energy consumption. [10] studied how the task request initiated by the business terminal is mapped to the edge node problem and the resource allocation problem to minimize the service delay, and finally proposes the BRT algorithm to solve the optimal task allocation and resource allocation decisions. [11] proposed a cooperative computing offloading framework based on three-tier mobile edge computing networks. At the same time, it also considered horizontal offloading between edge nodes. Joint

optimization of offloading decision-making and computing resource allocation is used to minimize average delay.

The route selection problem mainly studies the transmission route problem of tasks in the edge network. [12] studied the energy management problem of MEC on power-constrained equipment. In order to solve the trade-off between data processing capacity and energy efficiency, a sustainable strategic intelligence drives edge routing algorithm is proposed. [13] proposed a temporary on-demand multi-path distance vector routing protocol based on link lifetime and energy consumption prediction to save routing discovery when routing fails cost (Table 1).

3 Task Allocation Model Based on Multi-edge-node Cooperation

3.1 Business Requests and Resources

Edge nodes (EN) with computing and storage capabilities are deployed in the network, represented by set $\mathcal{M} = \{1, 2, \dots, M\}$, and there are a large number of user end (UE) accessing the network, represented by set $\mathcal{N} = \{1, 2, \dots, N\}$.

The task sent by UE i is represented by \mathcal{T}_i , which includes more than one subtasks represented by $w_{ij} = (c_{ij}, e_{ij}, t_{ij}, d_{ij}, \lambda_{ij})$. c_{ij} represents computing resource requirement, e_{ij} represents storage resource requirement, d_{ij} represents the amount of input data. t_{ij} represent the processing delay when the resource requirements are met. λ_{ij} represents the ratio of the result data amount to the input data amount. There is no temporal dependency between subtasks. \hat{t}_i represents the delay constraint of task \mathcal{T}_i .

Because of the resource heterogeneity of EN, containers and virtualization technologies are used to support the realization of EN resource allocation, and the amount of resources required by subtasks is represented by the number of virtual resource units. It is assumed that the task requests of all UEs are sent at the same time. The task allocation decision $X = \{x_{ijk}\}$ is specified as follows,

$$x_{ijk} = \begin{cases} 1 & \text{if subtask } j \text{ of UE } i \text{ is allocated to EN } k \\ 0 & \text{else} \end{cases} \quad (1)$$

Considering that each subtask must be calculated by one node, the following constraints need to be met:

$$\sum_{k \in \mathcal{M}} x_{ijk} = 1 \quad (2)$$

The resource of EN k is represented by (C_k, E_k) . C_k means the number of virtual computing resource units. E_k means the number of virtual storage resource units. Therefore, the task allocation decision X has the following constraints,

$$\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{T}_i} c_{ij} x_{ijk} \leq C_k, k \in \mathcal{M} \quad (3)$$

$$\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{T}_i} e_{ij} x_{ijk} \leq E_k, k \in \mathcal{M} \quad (4)$$

Table 1. Definition of variables

Symbol	Description
N	the number of UEs
M	the number of ENs
\mathcal{N}	UE set
\mathcal{M}	EN set
R_i	EN set for completing subtasks of UE i
\mathcal{T}_i	subtask set requested by UE i
w_{ij}	subtask j of UE i
c_{ij}	the computing resource requirement of w_{ij}
e_{ij}	the storage resource requirement of w_{ij}
d_{ij}	the amount of input data of w_{ij}
t_{ij}	the computing delay of w_{ij} if the resource requirement is met
\hat{t}_i	the delay constraint of \mathcal{T}_i
λ_{ij}	the ratio of the result data amount to the input data amount
C_k	the number of virtual computing units of EN k
E_k	the number of virtual storage unit of EN k
x_{ijk}	whether subtask j of UE i is allocated to EN k
$y_{i\bar{k}}$	whether the result of task \mathcal{T}_i is merged on EN \bar{k}
u_i	the access point of UE i
\mathcal{N}_k	the UE set associated with EN k
B_k	the bandwidth of EN k
p_i	the transmission power of UE i
h_{i,u_i}	the UE i and EN u_i channel gain
σ^2	the additive Gaussian white noise power
y_i	the signal-to-noise ratio of UE i and EN u_i
$f_{ij u_i k g}^{send}$	whether the g th routing path from u_i to k is selected for input data sending of w_{ij}
$f_{ijk \bar{k} g}^{merge}$	whether the g th routing path from k to \bar{k} is selected for result data merging of w_{ij}
$f_{i \bar{k} u_i g}^{return}$	whether the g th routing path from \bar{k} to u_i is selected for result data of returning of w_{ij}
$W_{kk'g}$	whether there is a g th path between EN k and EN k'
$H_{kk'g}$	number of hops of g th routing path between EN k and EN k'
$R_{kk'g}$	number of transmission bandwidth of g th routing path between EN k and EN k'
$T_{i,r_{i1},r_{i2}}^{finish}$	the cooperative completion delay of subtask \mathcal{T}_i by r_{i1}, r_{i2}
$T_{i,r_{i1},r_{i2}}^{comm}$	the data transmission delay of subtasks of UE i from EN r_{i1} to EN r_{i2}
$T_{i,r_{i1}}^{comp}$	the computing delay of the subtasks of UE i on EN r_{i1}

3.2 Delay

Each UE connect to the nearest EN. Assume the access point of UE i is EN $u_i \in \mathcal{M}$. The UE set associated with EN k can be represented as $\mathcal{N}_k = \{i : i \in \mathcal{N}, u_i = k\}$.

The bandwidth resource of EN k is B_k Hz. We defaults that the UEs associated with EN allocate the EN bandwidth resources evenly. The signal-to-noise ratio of UE i and EN u_i is $y_i = \frac{p_i h_{i,u_i}}{\sigma^2}$, where p_i represents UE i transmission power, h_{i,u_i} represents UE i and EN u_i channel gain, and σ^2 represents additive Gaussian white noise power. Therefore when the UE i accesses the EN u_i , the uplink data transmission rate of the UE i expressed as follow.

$$v_i = \frac{B_{u_i}}{|\mathcal{N}_{u_i}|} \log_2(1 + y_i) \quad (5)$$

The network delay from UE i to EN u_i includes the transmission delay caused by the port rate and propagation delay. c represent the propagation speed of the wired or wireless channel, and D_{i,u_i} represent the physical distance between the UE i and EN u_i . Therefore, the network delay of task input data transmission to access point EN can be expressed as,

$$T_{i,u_i}^{up} = \frac{\sum_{j \in \mathcal{I}_i} d_{ij}}{v_i} + \frac{D_{i,u_i}}{c} \quad (6)$$

Similar to [14, 15], the downlink bandwidth of the UE is much higher than the uplink bandwidth, and data size after task processing is usually much smaller than before, so we ignore the downlink transmission delay of sending the task results from EN to UE.

The computing task sent by the UE arrives at the computing node through multiple EN forwarding during transmission in the edge network. Assume that there are at most G routing paths available between any two EN, and $W_{kk'g}$ represent if there is a g th path between EN k and EN k' . Number of hops and transmission bandwidth of g th routing path between EN k and EN k' is represented by $H_{kk'g}$ and $R_{kk'g}$. For convenience, we assume that queuing delays on EN are not taken into account when data transferred over the network. We also assume that there will be no device or link failures in the network.

3.3 Data Route

The cooperative computing delay of task includes computing delay in edge node and data transmission delay in the edge network. Considering that the delay of task computation is fixed, it is only necessary to optimize the task allocation and route selection. The multi-edge-node cooperative computing process is shown in Fig. 2. Firstly, each subtask sent by UE starts from the access point and is transmitted to the corresponding computing node through the sending route, as shown in Fig. 2(a). After the subtask is completed, all the processing results of

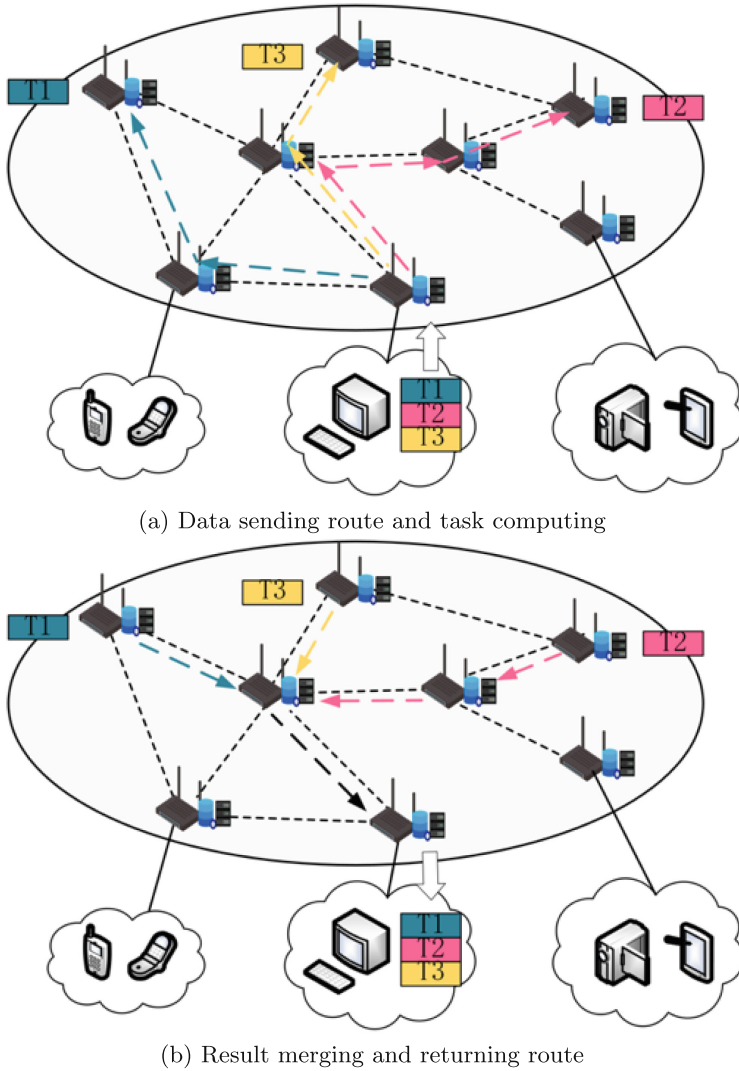


Fig. 2. The process of multi-edge-node cooperative computing

each subtask are sent to a EN to merge, and finally returned to the access point, as shown in Fig. 2(b).

According to the location of task data in the network, the route process in cooperative computing is divided into three stages: sending route, merging route, and returning route. The starting point of the sending route is the UE's access point, and the end point is the computing node, which is determined by the task allocation decision X . After the subtask is completed the start point of merging route is the computing node, and the end point is the merging node. The starting

point in the returning route is the merging node, and the end point is the access point. It can be seen that the selection of different merging nodes will have an impact on the delay of merging route and returning route. We set the merging node decision as $Y = \{y_{i\bar{k}}\}$, $i \in \mathcal{N}$, $\bar{k} \in \mathcal{M}$. Its value specification is as follows:

$$y_{i\bar{k}} = \begin{cases} 1 & \text{if merging node of task } \mathcal{T}_i \text{ is } \bar{k} \\ 0 & \text{else} \end{cases} \quad (7)$$

Since each task needs to select a merging node, there are the following constraints:

$$\sum_{\bar{k} \in \mathcal{M}} y_{i\bar{k}} = 1, i \in \mathcal{N} \quad (8)$$

We use $F^{send} = \{f_{iju_kg}^{send}\}$, $F^{merge} = \{f_{ijk\bar{k}g}^{merge}\}$, $F^{return} = \{f_{i\bar{k}u_i g}^{return}\}$ to represent the route selection strategy for the three stages. $f_{iju_kg}^{send}$ indicates whether to choose the g th path when the input data of subtask w_{ij} are transmitted to EN k in sending stage. $\{f_{ijk\bar{k}g}^{merge}\}$ and $\{f_{i\bar{k}u_i g}^{return}\}$ are the same as above. The value specification of $F^{send} = \{f_{iju_kg}^{send}\}$ is as follows:

$$f_{iju_kg}^{send} = \begin{cases} 1 & \text{The } g\text{th routing path from } u_i \text{ to } k \text{ is selected for} \\ & \text{data } w_{ij} \\ 0 & \text{else} \end{cases} \quad (9)$$

The value specification of F^{send} and F^{return} is the same as above.

Because the start point of the sending route is the UE's access point, the end point is determined by X , and only one path between them can be selected, so the constraints of F^{send} are as follows:

$$\sum_{k \in \mathcal{M}} (x_{ijk} \sum_{g=1}^G W_{u_kg} f_{iju_kg}^{send}) = 1, \forall i \in \mathcal{N}, j \in \mathcal{T}_i \quad (10)$$

$$\sum_{k \in \mathcal{M}} \sum_{g=1}^G f_{iju_kg}^{send} = 1, \forall i \in \mathcal{N}, j \in \mathcal{T}_i \quad (11)$$

In the same way, the start point of the merging route is determined by the task allocation decision X , and the end point is determined by Y , and only one path between them can be selected. So the constraints of F^{merge} are as follows:

$$\sum_{k \in \mathcal{M}} (x_{ijk} \sum_{\bar{k} \in \mathcal{M}} (y_{i\bar{k}} \sum_{g=1}^G W_{\bar{k}g} f_{ijk\bar{k}g}^{merge})) = 1, \forall i \in \mathcal{N}, j \in \mathcal{T}_i \quad (12)$$

$$\sum_{k \in \mathcal{M}} \sum_{\bar{k} \in \mathcal{M}} \sum_{g=1}^G f_{ijk\bar{k}g}^{merge} = 1, \forall i \in \mathcal{N}, j \in \mathcal{T}_i \quad (13)$$

The constraints of $F_{ij\bar{k}u_i,g}^{return}$ are as follows:

$$\sum_{\bar{k} \in \mathcal{M}} (y_{i\bar{k}} \sum_{g=1}^G W_{\bar{k}u_i,g} f_{i\bar{k}u_i,g}^{return}) = 1, \forall i \in \mathcal{N} \quad (14)$$

$$\sum_{\bar{k} \in \mathcal{M}} \sum_{g=1}^G f_{i\bar{k}u_i,g}^{return} = 1, \forall i \in \mathcal{N} \quad (15)$$

The delay calculation process of cooperative computing is as follows:

Sending Delay. The network transmission delay of the subtasks w_{ij} from the access point u_i to the computing node is:

$$T_{ij}^{send} = \sum_{k \in \mathcal{M}} \sum_{g=1}^G x_{ijk} W_{u_i k g} f_{ij u_i k g}^{send} H_{u_i k g} \frac{d_{ij}}{R_{u_i k g}}, \forall i \in \mathcal{N}, j \in \mathcal{T}_i \quad (16)$$

Computing Delay. Under the condition that the requirements of computing resource and storage resource are met, the computing delay of each subtask is $T_{ij}^{comp} = t_{ij}$

Result Merging Delay. After the subtask w_{ij} has been processed, the delay for sending the result to the merging node is:

$$T_{ij}^{merge} = \sum_{k \in \mathcal{M}} \sum_{\bar{k} \in \mathcal{M}} \sum_{g=1}^G x_{ijk} y_{i\bar{k}} W_{k\bar{k}g} f_{ijk\bar{k}g}^{merge} H_{k\bar{k}g} \frac{d_{ij}}{R_{k\bar{k}g}}, \quad (17)$$

$\forall i \in \mathcal{N}, j \in \mathcal{T}_i$

Result Returning Delay. After the results of \mathcal{T}_i all arrive at the merging node, the delay of data transmission to the access point u_i is:

$$T_i^{return} = \sum_{\bar{k} \in \mathcal{M}} \sum_{g=1}^G y_{i\bar{k}} W_{\bar{k}u_i,g} f_{i\bar{k}u_i,g}^{return} H_{\bar{k}u_i,g} \frac{\sum_{j \in \mathcal{T}_i} \lambda_{ij} d_{ij}}{R_{\bar{k}u_i,g}}, \forall i \in \mathcal{N}$$

Considering that the merging time depends on the latest subtask in \mathcal{T}_i that arrives at the merging node, the cooperative completion delay of \mathcal{T}_i is:

$$T_i^{coop} = \max_{j \in \mathcal{T}_i} \{T_{ij}^{send} + T_{ij}^{comp} + T_{ij}^{merge}\} + T_i^{return} \quad (18)$$

3.4 Problem Model

Combining Sects. 3.1, 3.2, and 3.3, the total delay T_i for task \mathcal{T}_i from sending request to receiving the result can be obtained as

$$T_i = T_{i,u_i}^{up} + T_i^{coop} + T_{i,u_i}^{down} \quad (19)$$

T_{i,u_i}^{down} represents the delay of the result from the EN u_i to the UE i . As described in Sect. 3.2, the delay is small and ignored. At the same time, the delay of generating task allocation and route selection decisions are ignored. Therefore, the task completion delay of the UE i is

$$T_i = T_{i,u_i}^{up} + T_i^{coop} \quad (20)$$

The average completion delay of all UE tasks is

$$\bar{T} = \frac{1}{N} \sum_{i \in \mathcal{N}} T_i \quad (21)$$

We describes task allocation and route selection as a problem of minimizing the average completion delay of tasks, as shown below:

$$\begin{aligned} \text{P1:} \quad & \min_{X, Y, F^{send}, F^{merge}, F^{return}} \bar{T} \\ \text{s.t. constraints} \quad & (1)-(4), (7)-(15) \end{aligned}$$

4 Algorithm

In this section, we propose an Improved Biogeography-based Optimization Algorithm in Task Allocation and Sending Route Optimization (IBBO-TASRO) to solve the problem of task allocation X and route selection F^{send} .

4.1 Task Allocation and Sending Route Strategy

Decisions in P1 includes task allocation X , three-stage route selection F^{send} , F^{merge} , F^{return} and merging node Y . It is difficult to solve for these variables at the same time. The solution space is relatively large and it is difficult to meet real-time requirements. Considering that in the case of meeting service resource requirements, the closer the node participating in the task computing and the result merging is to the access point, the shorter the data transmission path in the network, and the smaller the average task completion delay. Based on this idea, we initialize the merging node decision, set the merging node of each task at the access node EN u_i , $\bar{Y} = \{y_{ik} | \forall i \in \mathcal{N}, y_{iu_i} = 1, y_{ik} = 0, k \in \mathcal{M} \setminus \{u_i\}\}$. The delay of returning route is 0. We convert problem P1 to P2:

$$\begin{aligned} \text{P2:} \quad & \min_{X, F^{send}, F^{merge}} \bar{T} \\ & Y = \bar{Y} \\ \text{s.t. constraints} \quad & (1)-(4), (7)-(15) \end{aligned}$$

The IBBO-TASRO algorithm is used to solve P2. The IBBO-TASRO algorithm flow is as follows.

Algorithm 1. IBBO-TASRO**Require:** $L, S_{max}, I, E, m_{max}, \Omega, \mathcal{T}_i, i \in \mathcal{N}, C_k, E_k, k \in \mathcal{M}$.**Ensure:** X^*, F^{send*}

```

1: Initialize task allocation decisions  $X^l$  for each habitat.
2: for  $n = 1$  to  $\omega$  do
3:   for  $l = 1$  to  $L$  do
4:     for  $i = 1$  to  $N$  do
5:       for  $j$  to  $\mathcal{T}_i$  do
6:         calculate  $F^{send,l}, F^{merge,l}$  according to Eq.(X)
7:       end for
8:     end for
9:   end for
10:  calculate HIS according to  $\bar{T}(X, \tilde{Y}, F^{send}, F^{merge})$ 
11:  select the best top 2 habitat as collection  $\psi$ 
12:  for  $l = 1$  to  $L$  do
13:    if  $l \in \psi$  then
14:      continue
15:    end if
16:    if  $\text{rand}(0,1) < \lambda'(s_l)$  then
17:      for  $l' = 1$  to  $L$  do
18:        if  $\text{rand}(0,1) < \mu(s_l)$  then
19:          Generate a variable  $i$  randomly in range  $[0, D]$ 
20:           $x_i^l \leftarrow x_i^{l'}$ 
21:          if  $X^l$  doesn't meet constraint (3) or (4) then
22:            replace  $X^l$  with a new habitat within bounds.
23:          end if
24:        end if
25:      end for
26:    end if
27:  end for
28:  for  $l = 1$  to  $L$  do
29:    if  $l \in \psi$  then
30:      continue
31:    end if
32:    if  $\text{rand}(0, 1) < m(s_i)$  then
33:       $x_i^l \leftarrow 1 + \text{rand}(M)$ 
34:    end if
35:    if  $X^l$  doesn't meet C1 C2 then
36:      replace  $X^l$  with a new habitat within bounds.
37:    end if
38:  end for
39: end for

```

5 Experimental Results and Analysis

In order to verify the superiority of the joint optimization algorithm of task assignment and routing proposed in this paper, we conduct a simulation exper-

iment. The simulation environment was run on a 64 bit MacOS Sierra system with a 2.8 GHz Intel Core i7 processor and 16 GB of memory.

It is assumed that the simulation environment is a $1 \text{ km} \times 1 \text{ km}$ square area which contains 10 EN and 60 UE. EN and UE are randomly deployed in this area. The CPU frequency(GHz) and storage size(GB) of each edge node obey normal distribution, which are $\tilde{N}_1(10, 2 \times 10^2)$, $\tilde{N}_1(10^2, 3 \times 10^2)$. We set virtual computing unit to be 0.2 Ghz and virtual storage unit to be 0.6 GB. The number of subtasks of each UE follows a uniform distribution $U(2, 7)$. The requirements of virtual computing resource unit and virtual storage unit follow poisson distribution. The mean values of poisson distribution are $\lambda_1 = 9, \lambda_2 = 11, \lambda_3 = 50$. The data transmission rate(KB/s) between two edge nodes follows the normal distribution $\tilde{N}_3 = (3 \times 10^3, 10^3)$.

Figure 3 shows the convergence effect of the four algorithms. The comparison algorithms are Biogeography-based Optimization for Task Allocation and Sending Route Selection(BBO-TASRO), Genetic Algorithm for Task Allocation and Sending Route Selection(GA-TASRO) and Particle Swarm Optimization for Task Allocation and Sending Route Selection(PSO-TASRO). The four algorithms begin to converge after 100 iterations, and finally converge to 108, 122, 134, and 151 ms respectively at 500 iterations. It can be seen that IBBO-TASRO has the fastest convergence speed and the lowest average task completion delay. IBBO-TASRO speeds up the search rate and avoids the algorithm from falling into a local optimal solution through the elite retention strategy and the improved migration rate.

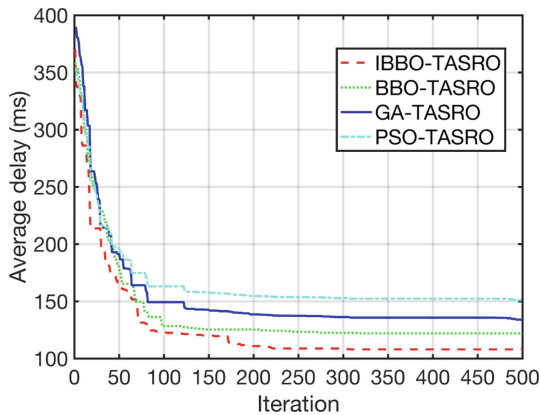


Fig. 3. Convergence performance of IBBO-TASRO and comparison with other algorithms

Figure 4 shows the average delay of the IBBO-TASRO under different UE numbers comparison with other algorithms. As the number of UE increases, the average delay of the four algorithms shows an upward trend. When the number of UE is 90, the average delays of the IBBO-TASRO, BBO-TASRO, PSO-TASRO,

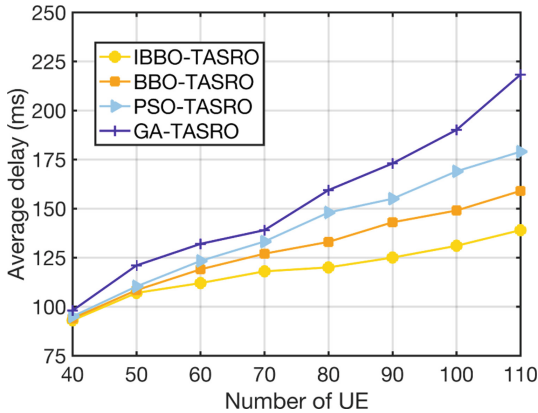


Fig. 4. Average delay of IBBO-TASRO with different number of UE

and GA-TASRO algorithm are 125, 146, 148, 173 respectively. Compared with BBO-TASRO, the delay of IBBO-TASRO is reduced by 1.4%. IBBO-TASRO uses the elite retention strategy in each algorithm iteration, avoiding the destruction of high-quality candidate solutions in the migration and mutation operation, and use the improved migration rate to enhance the stability of the algorithm convergence.

6 Conclusion

This paper proposes a cooperative computing mechanism based on edge routing optimization for power Internet of Things to minimize the service delay. First, a task allocation model based on multi-edge-node cooperation is established, and the model is transformed into the problem of minimizing the average task completion delay. Then the IBBO-TASRO algorithm is proposed to solve the task allocation and route selection. Finally the experimental results show that compared with other algorithms, the proposed algorithm has lower delay and can meet the resource requirements of a large number of business terminals at the same time.

References

1. Ou, Q., Zhen, Y., Li, X., et al.: Application of internet of things in smart grid power transmission. In: 2012 third FTRA International Conference on Mobile, Ubiquitous, and Intelligent Computing, pp. 96–100. IEEE (2012)
2. Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. *Comput. Netw.* **52**(12), 2292–2330 (2008)
3. Shi, W., Cao, J., Zhang, Q., et al.: Edge computing: vision and challenges. *IEEE Internet Things J.* **3**(5), 637–646 (2016)
4. Hayes, B.: Cloud computing (2008)

5. Gong, C., Lin, F., Gong, X., et al.: Intelligent cooperative edge computing in internet of things. *IEEE Internet Things J.* **7**(10), 9372–9382 (2020)
6. Tran, T.X., Hajisami, A., Pandey, P., et al.: Collaborative mobile edge computing in 5G networks: new paradigms, scenarios, and challenges. *IEEE Commun. Mag.* **55**(4), 54–61 (2017)
7. Zhu, S., Gui, L., Chen, J., Zhang, Q., Zhang, N.: Cooperative computation offloading for UAVs: a joint radio and computing resource allocation approach.: In: 2018 IEEE International Conference on Edge Computing (EDGE), San Francisco, CA, pp. 74–79 (2018). <https://doi.org/10.1109/EDGE.2018.00017>
8. Sun, P.J.: Privacy protection and data security in cloud computing: a survey, challenges, and solutions. *IEEE Access* **7**, 147420–147452 (2019)
9. Jararweh, Y., Doulat, A., AlQudah, O., et al.: The future of mobile cloud computing: integrating cloudlets and mobile edge computing. In: 2016 23rd International Conference on Telecommunications (ICT), pp. 1–5. IEEE (2016)
10. Niu, X., Shao, S., Xin, C., et al.: Workload allocation mechanism for minimum service delay in edge computing-based power Internet of Things. *IEEE Access* **7**, 83771–83784 (2019)
11. Wang, Y., Tao, X., Zhang, X., et al.: Cooperative task offloading in three-tier mobile computing networks: an ADMM framework. *IEEE Trans. Veh. Technol.* **68**(3), 2763–2776 (2019)
12. Trinh, H., Callyam, P., Chemodanov, D., et al.: Energy-aware mobile edge computing and routing for low-latency visual data processing. *IEEE Trans. Multimedia* **20**(10), 2562–2577 (2018)
13. Zhang, D.G., Chen, L., Zhang, J., et al.: A multi-path routing protocol based on link lifetime and energy consumption prediction for mobile edge computing. *IEEE Access* **8**, 69058–69071 (2020)
14. Du, J., Zhao, L., Feng, J., et al.: Computation offloading and resource allocation in mixed fog/cloud computing systems with min-max fairness guarantee. *IEEE Trans. Commun.* **66**(4), 1594–1608 (2017)
15. Chen, M., Hao, Y.: Task offloading for mobile edge computing in software defined ultra-dense network. *IEEE J. Sel. Areas Commun.* **36**(3), 587–597 (2018)



Voluntary Data Preservation Mechanism in Base Station-Less Sensor Networks

Yutian Chen¹, Jennifer Ly², and Bin Tang²(✉)

¹ Economics Department, California State University Long Beach, Long Beach, USA

Yutian.Chen@csulb.edu

² Computer Science Department, California State University Dominguez Hills,
Carson, USA

btang@csudh.edu

Abstract. We consider the problem of preserving a large amount of data generated inside *base station-less sensor networks* when sensor nodes are controlled by different authorities and behave selfishly. We modify the VCG mechanism to guarantee that each node, including the source nodes with overflow data packets, will voluntarily participate in data preservation. The mechanism ensures that each node truthfully reports its private type and network achieves efficiency for all the preserved data packets. Extensive simulations are conducted to further validate our results.

Keywords: VCG mechanism · Base station-less sensor networks · Data preservation

1 Introduction

Background. Wireless sensor networks are ad hoc multi-hop wireless networks formed by a large number of low-cost sensor nodes. They have been used in a wide range of applications such as military surveillance, environmental monitoring, and target tracking Yick et al. (2008). Recent years observe emerging wireless sensor networks deployed in challenging environments such as remote or inhospitable regions, or under extreme weather, to continuously collect large volumes of data for a long period of time. Such emerging sensor networks include underwater or ocean sensor networks Jang and Adib (2019); Ghaffarivardavagh et al. (2020), wind and solar harvesting Li et al. (2012); Cammarano et al. (2012), volcano eruption monitoring and glacial melting monitoring et al. (2020); Werner-Allen et al. (2006); Martinez et al. (2004), and seismic sensor networks Cochran et al. (2009). In these scenarios, it is not practical to deploy data-collecting base stations with power outlets in or near such inaccessible sensor fields. These sensor networks are referred to as *base station-less sensor networks (BSNs)* because of the absence of the base stations.

A major task of BSNs is that sensory data generated have to be stored inside the network for some unpredictable period of time before being collected by

periodic visits of robots or data mules Shah et al. (2003), or by low-rate satellite link Colitti et al. (2008). We focus on one scenario as follows. Some sensor nodes are close to the events of interest and are constantly generating sensory data, depleting their own storage spaces. These sensor nodes with depleted storage spaces (referred to as *source nodes*) need to offload their overflow data to sensor nodes with available storage (referred to as *storage nodes*) to avoid data loss. The process of preserving overflow data within the sensor network is called as *data preservation in base station-less sensor networks*.

As sensor nodes are characterized by limited battery power, storage spaces, and processing capacity, and wireless communication consumes most of the battery power of sensor nodes, the key challenge of data preservation in BSNs is to conserve sensors' battery power in this process. Tang et al. (2013) proposed a centralized algorithm to minimize the total energy consumption of data preservation in BSNs. They showed that this problem is equivalent to the minimum cost flow problem, which can be solved optimally and efficiently Ahuja et al. (1993). This centralized algorithm is applicable if all the storage nodes cooperate, in the sense that they willingly contribute their battery power and storage spaces to help to relay and store overflow data from source nodes.

Our Contributions. However, in large-scale distributed sensor networks, sensor nodes could be controlled by different users or authorities, each of which pursues self-interest. Therefore sensor nodes can behave selfishly only to maximize their own benefit. As data preservation unavoidably consumes sensor nodes' limited resources, including battery power and storage capacity, these sensor nodes have minimum or zero motivation to help with data preservation. In addition, data preservation cost parameters of sensor nodes can be specific to the nodes. For example, the type of battery of a sensor node could affect the energy it consumes in data preservation. Due to the complexity of related features among different sensor nodes, parameters of data preservation cost specific to a sensor node should be taken as its private information instead of treated as known by the public. While this cost information is fundamental for a centralized algorithm to configure efficient data preservation routes, selfish sensor nodes may not have an incentive to report their private cost parameters/types truthfully.

In the sight of issues rooted in selfishness, literature has adopted algorithmic mechanism design (AMD) Nisan (1999); Nisan and Ronen (1999; 2007), a subfield of game theory, to motivate selfish players to participate in the game (individual rationality) and also be truthful about their private type (efficiency). The central idea is to compensate (or charge) each player based on the task accomplished (or objects assigned) in the game. In this paper, we propose a voluntary data preservation mechanism by augmenting existing AMD techniques. In our model, each source node has a value of its overflow data and will choose to preserve the data only when its value exceeds the corresponding preservation cost. Therefore, our model avoids the inefficiency due to the *over preservation* of data, namely, when the data value falls short of its preservation cost but still gets preserved. In addition, our model requires source nodes to compensate storage nodes for help in preserving their data, thus shifting the compensation respon-

sibility from the central authority to the source nodes. Moreover, it guarantees that every single node in the network will *voluntarily* participate in the data preservation game, including the source nodes and the storage nodes. Therefore, by taking into consideration the data value, our model extends individual rationality to the entire network, and improves the network efficiency of data preservation. The central authority in our model serves mainly two functions: it calculates the efficient data preservation route of each data and the compensation of each node; it also serves as a residual claimant, who covers the deficit in the system when the amount paid by the source nodes is below the calculated compensation, or holds the surplus when the opposite is true.

Paper Organization. The rest of the paper is organized as follows. Section 2 reviews all the related work. Section 3 formulates the data preservation problem. Section 4 presents the algorithmic mechanism design approach and presents the voluntary VCG model. Section 5 presents our detailed simulation results and analysis. Section 6 concludes the paper with a discussion of future work.

2 Related Work

Data preservation in base station-less sensor networks has been studied extensively. It has been modeled as a minimum cost flow problem to achieve the energy optimization Tang et al. (2013); Crary et al. (2015) as well as fault-tolerance Tang et al. (2014). Xue et al. (2013) designed a maximum weighted flow algorithm to preserve data packets of different values. Some works studied a suite of new multiple traveling salesman placement problems for data aggregation Tang (2018); Tang et al. (2021). Recently, Hsu et al. (2020) designed a quadratic programming solution to maximize the survival time of preserved data packets for base station-less sensor networks. However, none of the above research considered the selfishness of sensor nodes.

Chen and Tang (2016) was one of the first to study data preservation in base station-less sensor networks while considering that sensor nodes are selfish. They took a game-theoretic approach and designed a computationally efficient and truthful Vickrey-Groves-Clark (VCG) mechanism Vickrey (1961); Groves (1973); Clarke (1971); Nisan (1999) wherein truth-telling is always a dominant strategy. Yu et al. (2022) further identified that when storage nodes have a limited amount of energy power, the VCG mechanism proposed in Chen and Tang (2016) is no longer truthful and nodes can manipulate the VCG mechanism in order to gain more utilities. They further designed a data preservation game that traces and punishes manipulative nodes in the BSN and delivers dominant strategies for truth-telling nodes while achieving provably optimal data preservation with cheat-proof guarantees.

However, there are several issues in the VCG model of Yu et al. (2022). First, they assume that all the overflow data will be stored in the network without considering the benefit and cost tradeoff of data preservation. As a result, no matter whether it is very valuable or of little value, each data will be preserved regardless of how much cost its preservation entails. Although total preservation cost

is minimized, the outcome may not be truly efficient for the system, as it could take much effort to preserve data with little value. Second, the central authority is in charge of making all the compensation payments to the system, which can be a considerable amount. Meanwhile, the source nodes, as the direct beneficiaries of data preservation, are not responsible for compensating any other node. Third, they assume that the source nodes will participate in data preservation and tell the truth, therefore do not need any motivation and compensation from the central algorithm. However, source nodes are also selfish as participating in data preservation is not costless to them.

Our work is inspired by Eidenbenz et al. (2005), which is a sender-centric truthful ad hoc routing protocol that considers the rational and selfish source nodes. However, in this work, the only private information of a sender is its willingness to pay to establish the connection with the destination, and the sender's action is simply to establish this connection or not. In contrast, in our work, the relationship between a node's cost parameters and its incurred costs is more complicated. By lying about different cost parameters to different extents, a node might manipulate its cost and switch from one action to another.

3 Data Preservation Problem Formulation

Network Model. The sensor network is represented as an undirected connected graph $G(V, E)$, where $V = \{1, 2, \dots, n\}$ is the set of n sensor nodes, and E is the set of edges. The sensory data are modeled as a sequence of data packets, each of which is a bits. Some sensor nodes are close to the event of interest and generate a large amount of data, which depletes their storage spaces. Nodes with overflow data are referred to as *source nodes*. There are nodes in the system with empty storage space for data storage, which are referred to as *storage nodes*. WLOG there are $k < n$ source nodes $V_s = \{1, 2, \dots, k\}$ and q storage nodes $V_{ss} = \{k + 1, k + 2, \dots, k + q\}$, with $k + q \leq n$. Thus the rest nodes viz. $V - V_s - V_{ss} = \{k + q + 1, k + q + 2, \dots, n\}$ have neither overflow data nor storage spaces and thus can only relay data from source nodes in the data preservation process. We refer to them as *transition nodes*. Note that any node, including the source node, storage node, and transition node, can carry out the function of relaying data.

Let d_i denote the number of overflow data source node i generates, which must be offloaded to some storage nodes to avoid being lost. Let $d = \sum_{i=1}^k d_i$ be the total number of overflow data, and let $D = \{1, 2, \dots, d\}$ denote the set of these d data, and j a single overflow data. Let $s(j) \in V_s$, $1 \leq j \leq d$, denote data j 's source node. Let $m_i > 0$ be the available free storage space (in bits) at storage node $i \in V_{ss}$. Note that $m_i = 0$ for any source or transition node i . We assume that $\sum_{i=1}^n m_i \geq d \cdot a$; that is, the total size of the overflow data is not larger than the total available storage spaces in the network. Moreover, all the overflow data has certain values, indicating the importance of the data in a

specific application. Let $g_i > 0$ be the value of each overflow data of source node $i \in V_s$.¹

We assume that each node has enough energy to participate in the data preservation process. Three different kinds of energy consumption are involved in data preservation, as described below:

- *Transmitting Energy* $E_i^t(j)$. When node i sends a data packet of a bits to its one-hop neighbor j over their distance $l_{i,j}$, the amount of *transmitting energy* spent by i is $E_i^t(j) = a \cdot \epsilon_i^a \cdot l_{i,j} + a \cdot \epsilon_i^e$. Here, ϵ_i^a is the energy consumption of sending one bit on transmit amplifier of node i , and ϵ_i^e is the energy consumption of transmitting one bit on the circuit of node i .
- *Receiving Energy* E_i^r . When node i receives an a -bit data packet from one of its one-hop neighbor, the amount of *receiving energy* it spends is $E_i^r = a \cdot \epsilon_i^e$. Here, ϵ_i^e is the energy consumption of receiving one bit on the circuit of node i . Note that E_i^r does not depend on the distance between nodes.
- *Storing Energy* E_i^s . When node i stores a -bit data into its local storage, the amount of *storing energy* it consumes is $E_i^s = a \cdot \epsilon_i^s$. Here ϵ_i^s is the energy consumption of storing one bit at node i .

Problem Formulation. A *preservation function* as $p : D \rightarrow V_{ss}$ indicates that a data packet $j \in D$ is offloaded from its source node $s(j) \in V_s$ to a storage node $p(j) \in V_{ss}$ to be preserved. Let $P_j = \{s(j), \dots, p(j)\}$ be the *preservation path* along which j is offloaded. Let $c_{i,j}$ denote node i 's energy consumption in preserving j , which be represented as Eq. (1) below, with $\sigma(i, j)$ being the successor node of i on P_j .

$$c_{i,j} = \begin{cases} E_i^t(\sigma(i, j)) & i = s(j) \\ E_i^r + E_i^s & i = p(j) \\ E_i^r + E_i^t(\sigma(i, j)) & i \in P_j - \{s(j), p(j)\} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The objective is to find a preservation function p and P_j ($1 \leq j \leq d$) to minimize the *total preservation cost*. The minimized total preservation cost is given by

$$c = \min_p \sum_{j=1}^d \sum_{i=1}^n c_{i,j} = \min_p \sum_{i=1}^n \sum_{j=1}^d c_{i,j}, \quad (2)$$

under the storage constraint that the total size of data offloaded to storage node i can not exceed i 's storage capacity: $|j|1 \leq j \leq d, p(j) = i \cdot a \leq m_i, \forall i \in V_{ss}$. The corresponding optimal preservation function is p^* and the optimal preservation path of data j is P_j^* .

Algorithm. Tang et al. (2013) has shown that this problem is equivalent to the minimum cost flow problem in a properly transformed graph of the sensor network graph. The minimum cost flow problem can be solved optimally

¹ For simplicity, we assume that all the overflow data at the same source node has the same value. Our model is readily extended to the case when a source node has different values over its overflow data.

and efficiently Ahuja et al. (1993). We adopt and implement the scaling push-relabel algorithm proposed in Goldberg (1997). It has the time complexity of $O(|V|^2 \cdot |E| \cdot \log(|V| \cdot C))$, where C is the maximum capacity of an edge in the transformed graph. We denote the algorithm designed in Tang et al. (2013) as *the centralized algorithm* to highlight that it minimizes data preservation energy based on the assumption that each node in the network is selfless and therefore fully cooperative.

Discussion. In this work, instead of cooperative sensor nodes, we consider the situation when each node is selfish, intended to maximize its own interest instead of the system's interest. We design a mechanism to incentivize selfish nodes to comply with efficient data preservation directed by the centralized algorithm. Note that incentives among different types of sensor nodes are diversified. First, as a source node $i \in V_s$ wants to offload its overflow data to be preserved, we assume that information related to i is public, including its cost parameters and value g_i of its overflow data. For a source node, it will have its data offloaded so long as the compensation it needs to pay to other helper nodes does not exceed the value it holds for the data. Second, a non-source node (i.e., a storage and transition node) has private information about its cost parameters, denoted as *private type*: $t_i = \{\epsilon_i^a, \epsilon_i^e, \epsilon_i^s\}, \forall i \in V - V_s$. The private types of each node need to be truthfully reported. On the other hand, the cost incurred by each node should be compensated, so that every node will willingly participate in data preservation.

Moreover, as we take into consideration the data value of each source node, the concept of efficiency includes two aspects. The first one is the *preservation efficiency* same as in the previous literature. That is, given the set of data to be offloaded, the total preservation cost in the system shall be minimized. This part can be solved using the centralized algorithm in Tang et al. (2013) under the condition that all the nodes willingly cooperate. The second one is referred to as *data efficiency*. If the data value is below the corresponding minimum preservation cost found by the centralized algorithm, the data should not be preserved. This is a novel feature in our work because it compares the benefit and cost of preserving each data, and thus truly maximizes the net gain from data preservation.

4 Voluntary VCG Mechanism

We first introduce the concepts and notations of the algorithmic mechanism design (AMD) model. The goal of the AMD model is to design a game in which selfish players who maximize their own utility have equilibrium strategies that result in a socially optimal outcome. If the equilibrium is a *dominant strategy solution*, we say that this mechanism is *strategy proof*.²

² Dominant strategy of a player is a strategy that maximizes his utility regardless of the other players' strategies. If each player has a dominant strategy, in equilibrium each will play his dominant strategy and the strategy profile is called the dominant strategy solution. Note that a dominant strategy solution is also a Nash equilibrium since no player has an incentive to unilaterally deviate from its strategy.

The AMD Model. There are n nodes in the network - node i has some private information t_i , called its type. There is an *output specification* that maps each type vector $t = \{t_1, \dots, t_n\}$ to some output o . Node i 's cost is given by *valuation function* $v_i(t_i, o)$, which depends on t_i as well as o . A *mechanism* defines for each node i is a set of strategies A_i . When i plays strategy $a_i \in A_i$, the mechanism computes an *output* $o = o(a_1, \dots, a_n)$ and a *payment vector* $p = (p_1, \dots, p_n)$, where $p_i = p_i(a_1, \dots, a_n)$. Node i wants to maximize its utility function $\pi_i(a_1, \dots, a_n) = v_i(t_i, o) + p_i$.

In the context of data preservation, for $i \in V - V_s$, the private type $t_i = \{\epsilon_i^e, \epsilon_i^a, \epsilon_i^s\}$. Node i 's strategy set A_i includes any value of private type t_i it can report. For $i \in V - V_s$, $v_i(t_i, o) = -c_{i,j}$ given by Eq. (1), and its utility is $\pi_i(t_i, t_{-i}) = p_i - \sum_{j=1}^d c_{i,j}$. I.e., a non-source node's utility is the difference between its received payment and its cost in data preservation. For a source node $i \in V_s$, our model will make it pay $g_i - c_{i,j}$ for each of its preserved data j , leaving node i zero utility from preserving its own data. On the other hand, a source node can relay other source nodes' data and will be compensated according to its cost for performing such tasks. Therefore, its utility is also zero by helping with other nodes' data preservation. Thus the payoff to a source node is zero.

As sensor nodes may have private information about the cost of data preservation, we can adopt the Vickrey-Groves-Clark (VCG) mechanism Vickrey (1961); Groves (1973); Clarke (1971), which is a strategy-proof mechanism Nisan (1999) frequently used when the objective function is the sum of all agents' valuations. Yu et al. (2022) used the VCG mechanism to motivate non-source nodes to participate in data preservation. Compared to Yu et al. (2022), our model will motivate all the nodes, including source nodes, to voluntarily participate in data preservation. As such, we name our mechanism as *voluntary VCG mechanism*. We start by formally defining the payment and utility of each node, denoted as the *payment model*.

4.1 Payment Model

We use c_i to denote node i 's true total cost in data preservation, and p_i the total payment to node i . Thus for each node $i \in V$, $\pi_i = p_i - c_i$. Let $t_{-i} = \{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n\}$ denote the vector of private types of all other nodes except node i , and $c_{-i} = \{c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_n\}$ denote the data preservation costs of all other nodes except node i .

Definition 1. Payment and Utility. *Based on Green and Laffont Green and Laffont (1979), under VCG mechanism, given any type \tilde{t}_i reported by node i , the amount of payment given to node $i \in V$ depends on whether node i is chosen to participate in data preservation according to the centralized algorithm. Its payment is 0 if it is not chosen, and its payment when it is chosen is:*

$$p_i(\tilde{t}_i, t_{-i}) = c_{V-\{i\}} - (\tilde{c}_V - \tilde{c}_i), \quad (3)$$

where $c_{V-\{i\}}$ is the minimum total cost of the preservation path when i is excluded from the network; \tilde{c}_V is the minimum total cost of the preservation

path when i is included in the network and reports its cost \tilde{t}_i . Therefore i 's utility is 0 when it is not chosen by the centralized algorithm; and when i is chosen, its utility is

$$\pi_i(\tilde{t}_i, t_{-i}) = p_i(\tilde{t}_i, t_{-i}) - c_i = c_{V-\{i\}} - (\tilde{c}_V - \tilde{c}_i) - c_i. \quad (4)$$

Moreover, we define c_V as the minimum total cost of the preservation path that goes through i when i truthfully reports its type, i.e., when $\tilde{t}_i = t_i$.

In addition, to decide whether to preserve data j or not requires knowing the associated payment due to the preservation of data j . We define the payment to storage node i due to its help in preserving data j as

$$p_{i,j}(\tilde{t}_i, t_{-i}) = c_{V-\{i\},j} - (\tilde{c}_{V,j} - \tilde{c}_{i,j}). \quad (5)$$

That is, based on the reported private types (\tilde{t}_i, t_{-i}) , the payment to node i for its help in preserving data j is the difference between the total data preservation cost of j when i is excluded and when i is involved, plus the reported preservation cost of j by node i . Here $c_{V,j}$ is the total preservation cost of j when the total preservation cost of the network is c_V ; and $c_{V-\{i\},j}$ is the total preservation cost of j given that i is removed therefore the total preservation cost of the network is $c_{V-\{i\}}$.

Let $h(i) = \{j \in D \mid i \in P_j, i \neq s(j)\}$ be the set of data for which node i is not its source node but belongs to its preservation path according to the centralized algorithm. That is, $h(i)$ is the set of data that node i helps (by either relaying or storing) in their preservation. Lemma 1 below says that node i 's payment $p_i(\tilde{t}_i, t_{-i})$ comes from two parts: one part is its help to data $j \in h(i)$ (called the direct help of i), and the second part is its help to data $j \notin h(i)$ (called as the indirect help of i). While the first part is straightforward, the second part is due to the holistic procedure of the centralized algorithm aiming at minimizing total preservation cost. Namely, removing node i may also affect the preservation path of those data which do not use node i for their preservation when node i is included in the network.

Lemma 1. *It holds that $p_i(\tilde{t}_i, t_{-i}) = \sum_{j \in D} p_{i,j}(\tilde{t}_i, t_{-i}), \forall i \in V - V_s$.*

Proof:

$$\begin{aligned} p_i(\tilde{t}_i, t_{-i}) &= c_{V-\{i\}} - (\tilde{c}_V - \tilde{c}_i) = \sum_{j \in h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j}] + \sum_{j \notin h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j}] + \tilde{c}_i \\ &= \sum_{j \in h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j}] + \sum_{j \notin h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j}] + \sum_{j \in h(i)} \tilde{c}_{i,j} \\ &= \sum_{j \in h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j} + \tilde{c}_{i,j}] + \sum_{j \notin h(i)} [c_{V-\{i\},j} - \tilde{c}_{V,j}] \\ &= \sum_{j \in D} p_{i,j}(\tilde{t}_i, t_{-i}). \end{aligned}$$

■

Time Complexity of the Payment Model. The time taken to compute the payment is the time taken for the minimum cost flow calculation, which is $O(|V|^2 \cdot |E| \cdot \log(|V| \cdot C))$, where C is the maximum capacity of an edge in the transformed graph Goldberg (1997).

An Example of Incentive to Lie. For a source node $s(j) \in V_s$ of data packet j , the incentive compatibility is that $s(j)$ pays no more than the value j if j is preserved. That is, to decide whether data j should be preserved or not, it compares $g_{s(j)}$ to $c_{V,j} = \sum_{i \in P_j} c_{i,j}$, the total preservation cost for data j , and preserves j only if $g_{s(j)} \geq c_{V,j}$. However, this could distort storage nodes' incentives and lead them to lie, as illustrated by the below example.

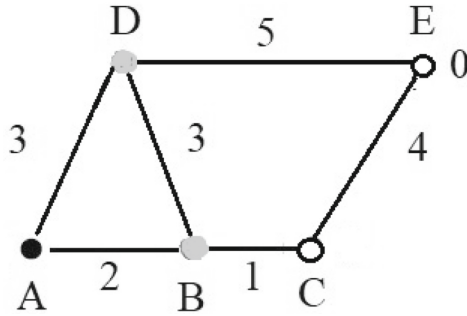


Fig. 1. An example for incentive to lie.

In Fig. 1, node A is the source node with 1 unit overflow data (data size is $a = 1$). Nodes C and E are storage nodes, each with 1 unit of storage capacity. Nodes B and D are transition nodes. The weights (i.e., the distance between the two end nodes) on edges AB , BC , AD , DE , DB , and EC are 2, 1, 3, 5, 3, 4, respectively. The cost parameters are $\epsilon_i^a = 1, \epsilon_i^e = 0$ for $i \in \{A, B, C, D, E\}$, and $\epsilon_C^s = 2, \epsilon_E^s = 0$. That is, the transmission cost is the same as the distance; the storage cost of C is 2 and of E is 0. Suppose the data value of node A is $g_A = 4$. Therefore, when each non-source node truthfully reports the private type, as the preservation path is A, B , and C , with a cost of $2 + 1 + 2 = 5 > 4$, the data will be dropped and $\pi_B = \pi_C = 0$.

Next, we show that each of B and C will have an incentive to lie to gain some positive utility. First, consider node B . Given that other nodes are telling the truth, let B lie by reporting $\epsilon_B^a = 0$. Then the cost along path A, B, C becomes $2 + 0 + 2 = 4 = g_A$, and data will be preserved. The utility of node B is $\tilde{\pi}_B = (3 + 5 + 0) - (2 + 0 + 2) + 0 - 1 = 3$, strictly higher than the case when it tells the truth (i.e., $\pi_B = 0$). Here, $3 + 5 + 0$ is the cost of the preservation path A, D, E , when B is removed from the network. Thus node B has an incentive to lie. Similarly, we can see that node C can be better off by lying to $\epsilon_C^s = 1$ when other nodes are truth-telling. Under lying, the path cost along ABC is $2 + 1 + 1 = 4 = g_A$ and the data will be preserved. Now

$\tilde{\pi}_C = (3 + 5 + 0) - (2 + 1 + 1) + 1 - 2 = 3$. Again $3 + 5 + 0$ is the path cost of preserving the data through D to E , when C is removed from the network. So node C has an incentive to lie.

Global Replacement Path. To preserve the incentives for storage nodes to report their types truthfully, we consider a methodology similar to in Eidenbenz et al. (2005) with several modifications. That is, to decide whether data j should be preserved, we compare j 's value to its *preservation cost along the global replacement path of j* . To find the global replacement path of j , we remove the original preservation path of j , i.e., P_j^* from the network, then let the central algorithm find the new path for j 's preservation, which is j 's global replacement path. The total preservation cost along this path is denoted as $c_{V-P_j^*,j}$. There are several diversifications from Eidenbenz et al. (2005). First, to remove the original path P_j^* , we only need to remove the non-source nodes on the path but not those source nodes (including j 's source node $s(j)$), since source nodes have no private information. Second, the removal includes the destination (the storage node) of j .³ We let a source node i preserve its data j as long as $g_{s(j)} \geq c_{V-P_j^*,j}$. If $g_{s(j)} \geq c_{V-P_j^*,j}$, we say that preservation of data j is feasible, and data j will be preserved. Otherwise, the preservation of data j is infeasible and data j will be dropped.

We define the set of data with feasible preservation as D^* . For D^* , the centralized algorithm will find the preservation path with a minimum total cost, which may or may not be the same as in P^* . As this will be the finalized preservation path used for the preservation of D^* , we denote it as P^f and the corresponding preservation path cost as c_V^f , with $c_V^f = \sum_{i \in P^f} \sum_{j \in D^*} c_{i,j}$. Correspondingly, $c_{V,j}^f = \sum_{i \in P_j^f} c_{i,j}$ is the total preservation cost of data j including its source node $s(j)$ and other nodes on the preservation path P_j^f . On the one side, with data j preserved, its value $g_{s(j)}$ is realized for $s(j)$, which will be paid by $s(j)$ to cover related cost. On the other side, according to the voluntary VCG mechanism, total payments for the preservation of j is $H_j \equiv \sum_{i \in V_s, i \in P_j} c_{i,j} + \sum_{i \in V-V_s} p_{i,j}$. Here $\sum_{i \in V_s} c_{i,j}$ is the total cost of the source nodes on the preservation path of data j (including j 's source node), as their costs in helping j 's preservation are directly observable; and $\sum_{i \in V-V_s} p_{i,j}$ is the total payment to all the other nodes due to their (direct or indirect) help in the preservation of data j . With some abuse of notation, we will call H_j as the *path payment of j* . Let the path payment along the final preservation path P^f be $H_j^f = \sum_{i \in V_s, i \in P_j^f} c_{i,j} + \sum_{i \in V-V_s} p_{i,j}^f$.

In general, for each data j , the payment made by the source node $g_{s(j)}$ and the finalized path payment H_j^f will not be equal, and the central authority will be the residual claimant to balance them. That is, if $g_{s(j)} > H_j^f$, there is a surplus in preserving data j , which will be held by the central authority; instead, if

³ In Eidenbenz et al. (2005), the removal of the preservation path does not include the destination, which is a common storage to every data. Also, the source does not participate in other nodes' data preservation in their work.

$g_{s(j)} < H_j^f$, there is a shortage in preserving data j , which will be covered by the central authority. Below we formally present the procedure of the voluntary VCG mechanism.

4.2 The Voluntary VCG Mechanism

Definition 2. The Voluntary VCG Mechanism. *It includes four stages:*

Stage 1. Each non-source node reports its private type t_i to the central authority.
Stage 2. For data set D , the centralized algorithm finds the optimal preservation path P^ , which minimizes the total preservation cost. For each data $j \in D$, it also finds the global replacement path and the corresponding replacement path cost $c_{V-P_j^*,j}$. The central authority then chooses to preserve data j if and only if $g_i \geq c_{V-P_j^*,j}$. The set of data chosen to be preserved is denoted as D^* .*
Stage 3. For D^ , the centralized algorithm finds the final optimal data preservation path P^f , and calculate payment to each node based on Eq. (3).*
Stage 4. Each of the nodes in the finalized data preservation path P^f chooses to participate in data preservation or not. For nodes who choose to participate, each source node i pays $g_i - c_{i,j}^f$ for each of its preserved data $j \in d_i, j \in D^$, and get reimbursed its cost in relaying other source nodes' data; each non-source node realizes the data preservation cost and also the payment given by Eq. (3), and gets utility given by Eq. (4).*

Note that each node makes strategic moves only in stages 1 and 4, whereas stages 2 and 3 are non-strategic: in the absence of base stations, the centralized algorithm is provided by an outsider of the system. Although the outsider is denoted as a central authority, it cannot enforce the outcome in the system: Each node makes decisions based on its own interest. Nonetheless, our major result (presented below) indicates that the voluntary VCG mechanism provides each node the incentive to truthfully report private type and also participate in data preservation as instructed by the central algorithm. Thus the mechanism achieves data preservation efficiency in the sense that it minimizes preservation cost for the set of preserved data. On the other side, there can be data dropped due to its exorbitant preservation cost (relative to its value), therefore our mechanism also improves data efficiency for the network.

Assumptions. Several assumptions are needed for the mechanism to work. First, the payment model and the algorithm are common knowledge to each node. Second, each node has enough energy to perform data preservation tasks. Third, the network satisfies the “minimum-energy 2-connectivity”, i.e., to any data, after a data preservation route is removed, there always exists an alternative route to preserve that data.⁴

⁴ This implies that $\sum_{i=k+1}^n m_{-i} \geq d \cdot a$. After removing any single storage node, the system can still preserve all the overflow data. Although the “minimum-energy 2-connectivity” requirement looks restrictive, Yu et al. (2022) found that it is satisfied with high probability, in general over 90% in their simulations.

For the voluntary VCG mechanism to be strategy-proof, it needs to satisfy two properties namely individual-rationality and incentive-compatibility, as explained below:

1. Individual-rationality (**IR**). It is the participation constraint that makes sure that each node, when truthfully reporting its type, will participate in data preservation once it is chosen by the centralized algorithm. That is, $\pi_i(t_i, t_{-i}) \geq 0 \forall t_{-i}$ and $\forall i \in V$.
2. Incentive-compatibility (**IC**). It requires that truthfully reporting private type is the dominant strategy of each node. Namely, each node gets the highest utility under truth-telling regardless of reported types of other nodes: $\pi_i(t_i, t_{-i}) \geq \pi_i(\tilde{t}_i, t_{-i}) \forall t_{-i}, \forall \tilde{t}_i \neq t_i$ and $\forall i \in V$.

Theorem 1. *The voluntary VCG mechanism satisfies the conditions IR and IC. In other words, it is a dominant strategy solution that each non-source node shall truthfully report its private type, and each node follows the centralized algorithm in data preservation.*

Proof: For a non-source node i that helps in preserving a data packet for which node i is not its source, node i either relays the data packet to a successor node, or stores the data packet. Therefore, it incurs one of the two costs below:

- *Relaying Cost* $c_i^r(j)$. When node i sends the data to one of its one-hop neighbor j over their distance $l_{i,j}$, its *relaying cost*, denoted as $c_i^r(j)$, is the sum of its receiving energy and transmitting energy. That is $c_i^r(j) = E_i^r + E_i^t(j) = 2 \cdot a \cdot \epsilon_i^e + a \cdot \epsilon_i^a \cdot l_{i,j}^2$.
- *Storing Cost* c_i^s . When node i receives a data packet and then stores it in its storage, its *storing cost*, denoted as c_i^s , is the sum of its receiving energy and its storing energy. That is, $c_i^s = a \cdot \epsilon_i^e + a \cdot \epsilon_i^s$.

For given D^* , the proof of IR and IC among non-source nodes is similar as in Yu et al. (2022) and is omitted here. On the other hand, the decision on D^* considers the replacement path cost of each data, therefore lying or not by a node does not affect D^* , leaving each node no incentive to lie.

For source nodes, they will pay g_i for each preserved data and get zero utility. Therefore, participating in data preservation is also a (weakly) dominant strategy of each source node. ■

5 Simulations

In this section, we conduct extensive simulations to validate our theoretical results. Our simulator is written in Python. For the minimum cost flow implementation, we use NetworkX net (2022), a Python package for network analysis. To visualize our theoretical analysis, we focus on 4×4 grid networks with 16 sensor nodes, with IDs from 0 to 15. Unless otherwise mentioned, in any of the 4×4 grid networks generated, five nodes are randomly selected as source nodes, each having one data item. The rest nodes are either storage nodes, each having

one or five units of storage spaces (i.e., $m = 1$ or 5); or transition nodes, each with zero storage spaces. Note that the number of storage nodes must be one more than the number of source nodes in order for the VCG mechanism to work. Each data point is the average of 20 simulation runs, and the error bars indicate 95% confidence intervals.

Energy Model. We adopt a simplified energy model as follows for the purpose of easy illustration, although our work is based on a more general energy model. We assume the energy consumption of either sending to receiving a data item is 0.5 and storing a data item costs zero. Consequently, the energy consumption of offloading one data item from any source node to any storage node is the number of hops the data item traverses in the grid. That is, the weight (or cost) of each edge is one unit of energy (0.5 of transmitting and 0.5 of receiving).

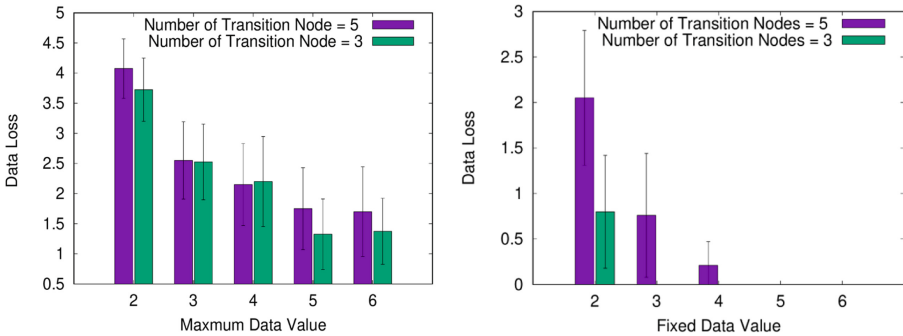


Fig. 2. The data loss in (a) random data values and (b) fixed data values. $m = 1$.

Investigating Data Loss. We first investigate the data loss in the network. Figure 2(a) shows the number of data losses by varying the values of the data items, which are random numbers between zero and a maximum data value. First, it shows that the number of data losses decreases with the increase of maximum data values. This is because our voluntary VCG mechanism preserves only the data items whose values exceed the costs of the corresponding replacement paths defined in Sect. 4.2. Therefore, the more valuable the data, the more likely it is preserved. Second, it shows that in most cases, the data loss for the number of transition nodes being 5 is larger than that when the number of transition nodes is 3. This is because more transition nodes generally result in a longer preservation path for a data item to reach its storage node. With a longer preservation path (and with a more significant cost), preserving data items becomes less desirable; thus more data loss occurs. In Fig. 2(b), we assign a fixed value to each data item and have the same observations as in Fig. 2(a), except that now the data loss for most of the data values are zeros for the number of transition nodes being 3. This is because being in the same range of $[2, 6]$, a fixed data value means more valuable data items than the random case in

Fig. 2(a). Consequently, both cases of transition numbers 5 and 3 have less data loss. In particular, for the number of transition nodes being 3, there are zero data losses for values ranging from 3 to 6; for the number of transition nodes being 5, there are zero losses for values at 5 and 6.

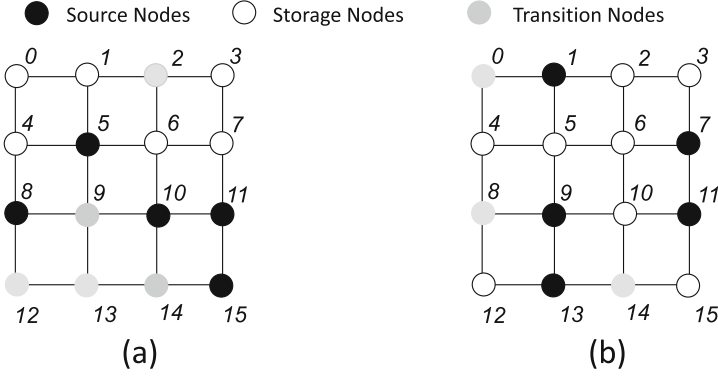


Fig. 3. The grid network (a) has 5 transition nodes while (b) has 3 transition nodes. Both have 5 source nodes and the rest are storage nodes.

Comparing Truth-Telling and Lying Utilities. Next, we validate the proposed VCG mechanism for the non-source nodes (i.e., storage and transition nodes) by comparing their truth-telling and lying utilities. We define the *scaling factor*, denoted as α , as the ratio between the reported and true values of a cost parameter of a non-source node. Following our energy model, when a non-source node reports (i.e., lies) about its energy cost of either transmitting or receiving one data item with a scaling factor of α , the reported costs become $0.5 \cdot \alpha$. When $\alpha < 1$, the node *under-reports* its cost by claiming it costs less energy than necessary; when $\alpha > 1$, it *over-reports* its cost by claiming it costs more energy than necessary; when $\alpha = 1$, it is truth-telling.

Table 1. Comparing truth-telling and lying utilities for non-source nodes with the number of transition nodes = 5. Node IDs with * are transitions nodes, the rest are storage nodes with $m = 5$.

Non-source node	0	1	2*	3	4	6	7	9*	12*	13*	14*
Truth-telling (i.e., $\alpha = 1$)	0.0	0.0	0.0	0.0	2.0	1.0	2.0	0.0	0.0	0.0	0.0
Over-reporting, $\alpha = 2$	0.0	0.0	0.0	0.0	2.0	1.0	2.0	0.0	0.0	0.0	0.0
Under-reporting, $\alpha = 0.5$	0.0	0.0	0.0	0.0	2.0	1.0	2.0	0.0	0.0	0.0	0.0

Number of Transition Nodes = 5. Figure 3(a) shows a randomly generated grid with nodes 5, 8, 10, 11, and 15 being source nodes, each having one data item.

The non-source nodes include 5 transition nodes 2, 9, 12, 13, and 14 and six storage nodes 0, 1, 3, 4, 6, and 7. We first consider that $m = 5$, i.e., each storage node can store all the data items, as shown in Table 1. The three rows indicate a non-source node’s utility under truth-telling (i.e., $\alpha = 1$), over-reporting with $\alpha = 2$, and under-reporting with $\alpha = 0.5$, respectively. First, each non-source node’s lying utility is no more than its truth-telling utility, indicating that truth-telling is the dominant strategy for the non-source nodes. Second, it shows the truth-telling utilities of most non-source nodes are zeros. This is because as each storage node has a storage capacity of 5, each can store all the data items in the network. As such, each of them is not critical as removing any of them does not increase the total preservation cost in the network. Third, nodes 4, 6, and 7 have the same utility under truth-telling and lying. This is because under either truth-telling or lying, the data preservation paths for the data items are mainly the same, due to regular topologies of the grid network.

Table 2. Comparing truth-telling and lying utilities for non-source nodes with number of transition nodes = 5. Node IDs with * are transitions nodes, the rest are storage nodes with $m = 1$.

Non-source node	0	1	2*	3	4	6	7	9*	12*	13*	14*
Truth-telling (i.e., $\alpha = 1$)	0.0	1.0	0.0	3.0	3.0	3.0	6.0	0.0	0.0	0.0	0.0
Over-reporting, $\alpha = 2$	0.0	1.0	0.0	3.0	3.0	3.0	6.0	0.0	0.0	0.0	0.0
Under-reporting, $\alpha = 0.5$	0.0	1.0	0.0	3.0	3.0	3.0	6.0	0.0	0.0	0.0	0.0

We then consider that each storage node can only store one data item in the BSN (i.e., $m = 1$), as shown in Table 2. Again, we observe that truth-telling utility is the dominant strategy for the non-source nodes. Compared to Table 1, it shows non-source nodes have more positive truth-telling utilities. This is because as each storage node has a storage capacity of 1, removing any of them could prolong the data items’ preservation paths, resulting in a higher total preservation cost for the entire network and a positive marginal cost for each node.

Number of Transition Nodes = 3. Finally, we consider that the number of transition nodes is 3. We found that only one node, storage node 12, has positive utilities. This is because 12 is critical to source node 13’s data preservation; however, other storage nodes are not critical to any source nodes’ data preservation. Due to the specific layout of the nodes, $m = 5$ or $m = 1$ do not make any difference in the utility computation (Table 3).

Table 3. Comparing truth-telling and lying utilities for non-source nodes with the number of transition nodes = 3. Node IDs with * are transitions nodes, the rest are storage nodes with $m = 5$ or $m = 1$.

Non-source node	0*	2	3	4	5	6	8*	10	12	14*	15
Truth-telling (i.e., $\alpha = 1$)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
Over-reporting, $\alpha = 2$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
Under-reporting, $\alpha = 0.5$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0

6 Conclusion and Future Work

In this work, we study the data preservation problem in base station-less sensor networks wherein sensor nodes behave selfishly. Selfishness is reflected in two aspects. First, for non-source nodes, their cost parameters are private information. They may not want to participate in data preservation nor have the incentive to report their private information truthfully. Second, source nodes may not have an incentive to store data once the data value exceeds the payment needed for its preservation. We design a voluntary VCG mechanism under which the individual sensor nodes, motivated solely by self-interest, achieve a good system-wide data preservation solution. In particular, the mechanism guarantees truthfulness among non-source nodes and data preservation efficiency among the data set chosen to be preserved. In addition, the mechanism makes each source node pay only its data value to preserve the data, thus guaranteeing the voluntariness of the source nodes.

Currently, we adopt grid topologies for BSNs for ease of illustration and visualization. In the future, we will consider a more realistic BSN topology wherein energy consumption of data preservation depends on the distance among nodes. We will also investigate the budget imbalance of the voluntary VCG model when storage nodes are energy-constrained. After that, we will validate theoretical findings using simulation results. First, simulation results shall verify the truthfulness and efficiency of the mechanism by contrasting the utility of each non-source node under the truth-telling strategy to what it is under lying. Second, the simulation will illustrate the number of data dropped due to their considerable preservation cost, and examine how such number changes in the network topology. Third, the simulation shall look into the budget imbalance of the voluntary VCG. While it verifies the upper-bound budget imbalance without capacity constraint, the simulation shall study the budget imbalance in the scenario with capacity constraint. Other future work includes extending our analysis to a dynamic scenario wherein overflow data are generated from time to time at different nodes. It is well understood in game theory that an infinitely repeated game gives a much larger set of equilibrium and in certain scenarios full cooperation can be achieved. In our setting of data preservation among selfish nodes, it is interesting to see to what extent we need to provide motivation for selfish nodes to cooperate and achieve optimal data preservation.

Acknowledgment. This work was supported in part by NSF Grants CNS-2131309, CNS-1911191, and CNS-1923956.

References

- NetworkX: Network Analysis in Python (2022). <https://networkx.org/>
- Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network Flows: Theory, Algorithms, and Applications. Prentice-Hall Inc., Hoboken (1993)
- Cammarano, A., Petrioli, C., Spenza, D.: Pro-energy: a novel energy prediction model for solar and wind energy-harvesting wireless sensor networks. In: IEEE 9th International Conference on Mobile Adhoc and Sensor Systems (MASS) (2012)
- Chen, Y., Tang, B.: Data preservation in base station-less sensor networks: a game theoretic approach. In: Proceedings of the 6th EAI International Conference on Game Theory for Networks (GameNets 2016) (2016)
- Clarke, E.H.: Multipart pricing of public goods. *Public Choice* **11**, 17–33 (1971)
- Cochran, E., Lawrence, J., Christensen, C., Chung, A.: A novel strong-motion seismic network for community participation in earthquake monitoring. *IEEE Inst. Meas.* **12**(6), 8–15 (2009)
- Colitti, W., Steenhaut, K., Descouvemont, N., Dunkels, A.: Satellite based wireless sensor networks: global scale sensing with nano- and pico-satellites. In: Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys 2008), pp. 445–446 (2008)
- Crary, N., Tang, B., Taase, S.: Data preservation in data-intensive sensor networks with spatial correlation. In: Proceedings of the International Workshop on Mobile Big Data (MobiData 2015) in Conjunction with Mobihoc 2015 (2015)
- Eidenbenz, S., Resta, G., Santi, P.: COMMIT: a sender-centric truthful and energy-efficient routing protocol for ad hoc networks with selfish nodes. In: 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS) (2005)
- Terray, L., et al.: From sensor to cloud: an IoT network of radon outdoor probes to monitor active volcanoes. *Sensors* **20**, 10 (2020)
- Ghaffarivardavagh, R., Afzal, S.S., Rodriguez, O., Adib, F.: Ultra-wideband underwater backscatter via piezoelectric metamaterials. In: Proceedings of the ACM SIGCOMM (2020)
- Goldberg, A.V.: An efficient implementation of a scaling minimum-cost flow algorithm. *J. Algorithms* **22**(1997), 1–29 (1997)
- Green, J., Laffont, J.: Incentives in public decision making. *Stud. Public Econ.* **1**(1979), 65–78 (1979)
- Groves, T.: Incentives in teams. *Econometrica* **41**, 617–631 (1973)
- Hsu, S., Yu, Y., Tang, B.: *DRE*²: achieving data resilience in wireless sensor networks: a quadratic programming approach. In: Proceedings of IEEE MASS (2020)
- Jang, J., Adib, F.: Underwater backscatter networking. In: Proceedings of the ACM SIGCOMM (2019)
- Li, Y., Li, X., Wang, P.: A module harvesting wind and solar energy for wireless sensor node. In: Wang, R., Xiao, F. (eds.) CWSN 2012. CCIS, vol. 334, pp. 217–224. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36252-1_20
- Martinez, K., Ong, R., Hart, J.K.: Glacisweb: a sensor network for hostile environments. In: Proceedings of SECON (2004)
- Nisan, N.: Algorithms for selfish agents. In: Meinel, C., Tison, S. (eds.) STACS 1999. LNCS, vol. 1563, pp. 1–15. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-49116-3_1

- Nisan, N., Ronen, A.: Algorithmic mechanism design. In: Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing (STOC 1999), pp. 129–140 (1999)
- Nisan, N., Ronen, A.: Algorithmic mechanism design. *Games Econ. Behav.* **35**(2007), 166–196 (2007)
- Shah, R.C., Roy, S., Jain, S., Brunette, W.: Data mules: modeling a three-tier architecture for sparse sensor networks. In: Proceedings of SNPA (2003)
- Tang, B.: *DAO²*: overcoming overall storage overflow in intermittently connected sensor networks. In: Proceedings of IEEE INFOCOM 2018 (2018)
- Tang, B., Jaggi, N., Takahashi, M.: Achieving data K-availability in intermittently connected sensor networks. In: Proceedings of the International Conference on Computer Communications and Networks (ICCCN) (2014)
- Tang, B., Jaggi, N., Wu, H., Kurkal, R.: Energy-efficient data redistribution in sensor networks. *ACM Trans. Sensor Netw.* **9**(2), 11:1–11:28 (2013)
- Tang, B., Ngo, H., Ma, Y., Alhakami, B.: *DAO²*: overcoming overall storage overflow in intermittently connected sensor networks. Technical report 2021-1. http://csc.csudh.edu/btang/papers/infocom18_journal.pdf. Computer Science Department, CSUDH
- Vickrey, W.: Counterspeculation, auctions and competitive sealed tenders. *J. Finance* **16**, 8–37 (1961)
- Werner-Allen, G., Lorincz, K., Johnson, J., Lees, J., Welsh, M.: Fidelity and yield in a volcano monitoring sensor network. In: Proceedings of OSDI (2006)
- Xue, X., Hou, X., Tang, B., Bagai, R.: Data preservation in intermittently connected sensor networks with data priorities. In: Proceedings of SECON (2013)
- Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. *Comput. Netw.* **52**(12), 2292–2330 (2008)
- Yu, Y., Hsu, S., Chen, A., Chen, Y., Tang, B.: Truthful and efficient data preservation in base station-less sensor networks. Technical report 2022-1 (2022). http://csc.csudh.edu/btang/papers/bsn_new.pdf. Computer Science Department, CSUDH. Submitted to ACM Transactions on Sensor Networks



Design and Implementation of Targeted Poverty Alleviation System Based on Blockchain Network

Jun Zhao^(✉), Wei Liang, Jingwen Li, Yan Liu, and Yan Liang

China Telecom Research Institute, Beijing, China
zhaojun8@chinatelecom.cn

Abstract. Targeted poverty alleviation is an important strategy to win the battle against poverty. In order to solve the problems of capital misappropriation and false poverty alleviation in the process of targeted poverty alleviation, a targeted poverty alleviation system based on blockchain network was designed and implemented. The system creatively draws on the characteristics and advantages of consensus mechanism, irreducible modification, traceability, distributed ledger and decentralization in blockchain network to the practice of precision poverty alleviation, and realizes “blockchain + targeted poverty alleviation” rural service integration platform. The system is applied to the digital supervision of poverty alleviation funds from village-to-household in the promotion of targeted poverty alleviation work by poverty alleviation units at all levels, and truly realizes that the poverty alleviation data in each link is true, reliable, provable, traceable and auditable. At the same time, it provides solid underlying credible data support for the “last mile” problem faced by village-to-household in targeted poverty alleviation work.

Keywords: Blockchain network · Precision poverty alleviation · Non-tampering · Traceability

1 Introduction

The report of the 19th National Congress of the Communist Party of China pointed out that “from now to 2020, it is a decisive period for building a moderately prosperous society in an all-round way”, and it emphasized the need to “resolutely win the tough battle against poverty [1]”. At present, all localities and departments are implementing the basic strategy of targeted poverty alleviation [2]; in order to win the poverty alleviation war, they are committed to innovating work models and overcoming difficulties. While my country has made great achievements in poverty alleviation, there are also many problems. At present, the main problems faced by targeted poverty alleviation work are as follows: insufficient information sharing between relevant governments and poverty alleviation objects, resulting in false poverty alleviation [3] or poor poverty alleviation work; the opaque use of poverty alleviation funds makes the process of poverty alleviation inconvenient to supervise [4, 5], and the problem of power being commodified easily

occurs; the lack of a system to integrate information from all parties leads to slow updating of data and information, and information exchange is limited; the identification of the poor is “only out but not in”, and there is a lack of support for people who have been lifted out of poverty. Dynamic tracking and precise management [6].

With the continuous development of Internet technology, the advantages of blockchain technology [7] have brought new ideas for targeted poverty alleviation and provided technical support. The traceability mechanism and time stamp mechanism [8] in the blockchain ensure the authenticity of the data in the targeted poverty alleviation work, and also solve the “opaque” and “inaccurate” problems in the targeted poverty alleviation work; the smart contract mechanism [9] is conducive to the combination of approval, service and supervision [10]. Therefore, it is necessary to apply blockchain technology to targeted poverty alleviation to ensure that poverty alleviation data cannot be tampered with and the source of poverty alleviation funds can be traced back.

This paper proposes to design and implement a provable, traceable and auditable targeted poverty alleviation system using blockchain technology. The contributions of our research are as follows:

1. This research establishes an authentic and credible basic data management system, and deploys a blockchain platform architecture with the advantages of data transparency, distributed storage, and trustworthiness for targeted poverty alleviation projects. Our research promotes the establishment of a tamper-proof and traceable information integration mechanism in targeted poverty alleviation projects.
2. A safe and efficient data exchange and sharing platform have been built to ensure the efficiency and safety of the information flow of targeted poverty alleviation projects, and to achieve traceable data usage. The platform provides a channel for data sharing of targeted poverty alleviation projects, thereby expanding service models and strengthening multi-departmental collaboration capabilities.
3. Created a regulatory audit service center for intelligent analysis. The center is committed to comprehensively controlling the progress of the implementation of targeted poverty alleviation projects and the use of funds, supporting dynamic performance assessment; improving data analysis capabilities and processing efficiency, and providing accurate, effective, and reliable data support for poverty alleviation decision-making.

2 System Design and Implementation

2.1 System Architecture Design

The targeted poverty alleviation system proposed in this paper involves three layers, namely the user layer, the application layer and the blockchain layer. Its overall design is shown in Fig. 1.

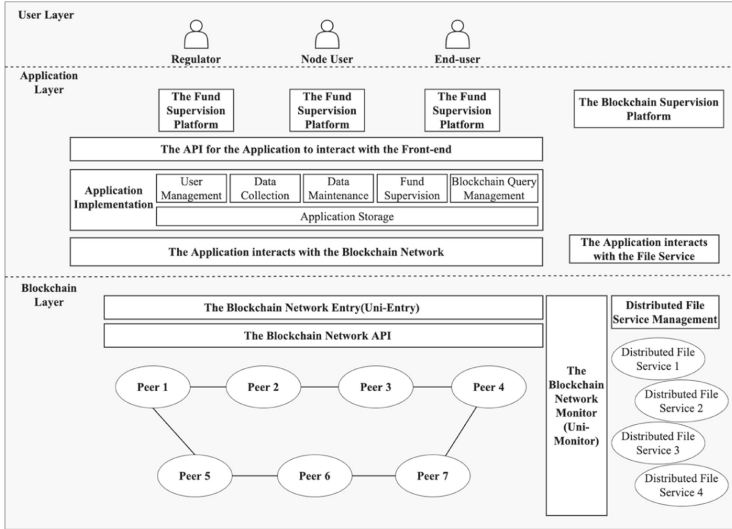


Fig. 1. System structure

2.2 User Layer

The user layer mainly involves three types of user roles: system management roles, system node roles and common user roles. The different user roles are defined as follows:

1. The system management user is responsible for the distribution of node user rights and the operation, management and maintenance of the blockchain system.
2. The blockchain node user is responsible for data entry and maintenance during the tracking process.
3. The end-user of the blockchain system can query the whole process information of traceability and anti-counterfeiting through the web platform.

2.3 Application Layer

Poverty Alleviation Fund Supervision Traceability Platform. The fund supervision traceability platform provides corresponding functions for different user roles, including user management module, data collection and maintenance module, blockchain query module, etc.

The user management module includes functions such as user account assignment and maintenance, and user authority assignment and maintenance.

The module of data collection and maintenance includes data collection function and maintenance function. The collected data is temporarily stored in the database of the application layer, and then stored in the block of the blockchain after the node confirms it.

The blockchain management query module includes functions such as blockchain management, blockchain data statistics, blockchain data query, etc. All query requests of this module are for data on the blockchain.

Poverty Alleviation Fund Supervision Traceability Query Platform. The poverty alleviation fund traceability query platform can provide various users with the ability to query the traceability data of poverty alleviation funds.

This platform mainly obtains the tracking information of poverty alleviation funds in each link by calling the query API interface provided by the blockchain. The platform is user-friendly. The traceability tracking data is displayed in the way of traceability, including the tracking data of each link, the collected documents and picture data, etc.

Blockchain Monitoring Platform. The blockchain monitoring platform provides blockchain network monitoring capabilities for blockchain system management roles.

The platform monitors the operation indicators of each node of the blockchain in real time. When the indicator of a node in the blockchain network exceeds the threshold, it will judge the node as abnormal and send an alarm to the relevant personnel by email, so that they can maintain the stable operation of the system in time.

2.4 Blockchain Layer

Blockchain Node. The blockchain network deploys 7 nodes, respectively deploying blockchain nodes on business stakeholders to form a blockchain node cluster. Among them, 4 nodes are deployed by the regulatory department, and 3 nodes are deployed by other institutions.

Blockchain Storage. The information involved in traceability tracking has various data forms, including metadata description form, image form, document form, etc.

The system will further sort out the collected data fields, classify the data fields according to different business links, and design them into a standard metadata description structure. The metadata description information is entered into the chain in JSON structure and stored in the block of the blockchain. The relevant pictures and documents collected during the process will be stored on a dedicated file server, and the files themselves will not enter the chain. Instead, the hash calculation will be performed on the picture and document file directories, and the unique identification of the files obtained after the calculation will be uploaded to the block.

Blockchain Entry. The blockchain provides creation and query interfaces to external business requests through APIs. The application layer cannot directly access the blockchain node machine, but access the blockchain through the unified entry of the blockchain. At the same time, the blockchain entry provides functions such as request load balancing, illegal request filtering, and high service availability.

Blockchain API Interface. The blockchain network will provide two types of APIs, data creation and data query.

The data creation API realizes the in-chain operation of collecting metadata in each link; the data query API realizes the query operation of the data on the chain, such as the query of the collected data of each link, and the query of the whole-process traceability data.

Blockchain Monitoring. A monitoring data collection module is deployed at each node of the blockchain to collect business indicator data of each link such as hardware resource occupancy and consensus on building blocks of the blockchain, providing monitoring data sources for the blockchain monitoring platform.

Blockchain Deployment. The deployment of the blockchain is shown in Fig. 2.

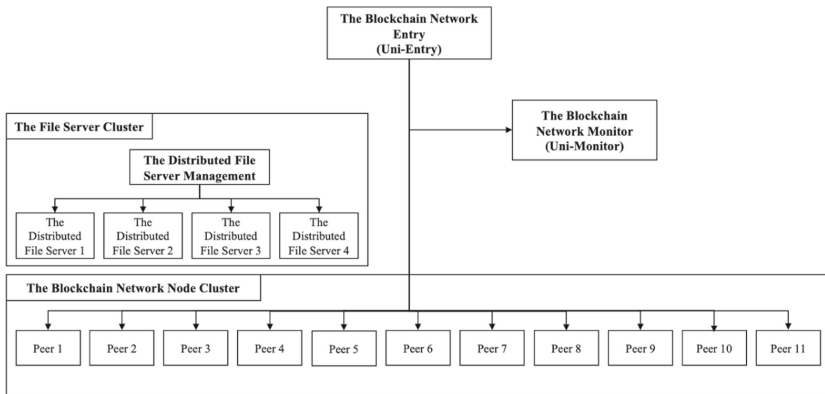


Fig. 2. The deployment of the blockchain

The occupancy of deployment machines is shown in Table 1 and Table 2.

Table 1. Block chain deployment machine occupancy.

Application deployment	Occupation of machines and uses
Uni-Entry	2 servers (nginx master/standby service)
Uni-Monitor	1 server (Grafana + InfluxDB, etc.)
Uni-Chain	7 servers (Unichain + RDb + LDb + collected, etc.) (4 nodes for regulatory authorities, 3 nodes for other related parties)

Table 2. File service deployment machine occupancy.

Application deployment	Occupation of machines and uses
Distributed file service management node	1 server
Distributed file service storage node	4 servers

3 The Practice of Blockchain in the Process of Targeted Poverty Alleviation

This system is mainly used in the informatization supervision of village-to-house poverty alleviation funds in the promotion of targeted poverty alleviation by poverty alleviation units at all levels. For various types of funds such as central financial funds, provincial financial funds, county financial funds, and industry poverty alleviation funds, it can achieve non-tamperable ledger operations and list management, and conduct inquiries and statistical analysis of all poverty alleviation funds.

The regulatory authorities can enter the information involved in the poverty alleviation fund policy, account arrival, allocation, integration, and other links into the blockchain system through the blockchain node. This information includes land and document number, direction of funding support, year of fund allocation, fund type, project to which the fund belongs, fund disbursement unit, fund receiving unit, fund amount, fund disbursement batches, etc. System users, such as fund sources, regulatory authorities, and poverty alleviation targets, can query and view traceable fund flow information on the platform at any time, as shown in Fig. 3.

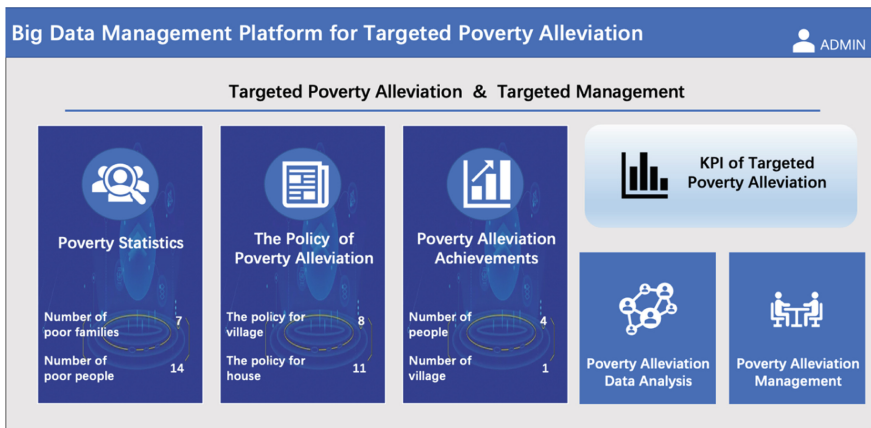


Fig. 3. The precision poverty alleviation big data platform

At present, the targeted poverty alleviation big data platform has covered 12 regions including Gansu, Yunnan, Shaanxi, Jiangxi, Henan, Hainan, Xinjiang, and Qinghai, serving more than 40% of the poor in China.

4 Conclusion

The application of blockchain technology to targeted poverty alleviation work provides a unified window for solving poverty alleviation and livelihood problems for the poor at the grassroots level. Blockchain technology enables tamper-proof poverty alleviation data and traceability of poverty alleviation funds, ensuring that the grassroots can effectively

enjoy preferential policies. It provides a convenient channel for the people to feedback problems, social supervision and evaluation, and the government to solve problems. This research provides a feasible solution for really promoting the broad participation and deep participation of the masses in poverty alleviation work. At the same time, it realizes targeted and meticulous work mode and two-way feedback for the whole process of poverty alleviation, such as targeted poverty situation identification, targeted policy implementation, and targeted poverty status cancellation. The proposed model ensures that the effectiveness of poverty alleviation is recognized by the people.

The application of blockchain technology in the field of targeted poverty alleviation has great advantages. Relying on the blockchain can solve the problems of information asymmetry and illegal scale, and achieve low-cost and high-efficiency coverage of poverty alleviation funds.

References

1. Wang, T.: Research on Xi Jinping's thought of the Fight Against Poverty in New Era. Jilin University (2018)
2. Lei, M.: On Xi Jinping's strategic thought for poverty alleviation. *J. Nanjing Agric. Univ. (Soc. Sci.)* **18**(01), 1–11+160 (2018)
3. Zhao, Z.: A study on the countermeasures of targeted poverty alleviation in the view of building a moderately prosperous society in all respects. Xi'an University of Technology (2018)
4. Hubei, X.: Municipal Bureau of Finance, Fan, J., Ma, S., Ge, W.: Research on the supervision of project funds in the field of poverty alleviation - taking Gucheng County of Xiangyang City, Hubei Province as an example. *Fiscal Science* **01**, 130–137 (2019)
5. Su, P.: Research on the supervision of the national rural poverty alleviation fund. Zhengzhou University (2017)
6. Ma, Y.: Analysis of the current situation of targeted poverty alleviation management in the new era. *Spec. Econ. Zone* **07**, 71–73 (2019)
7. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system. *Decentralized Bus. Rev.* 21260 (2008)
8. Zhang, L., Liu, B.X., Zhang, R.Y., Jiang, B.X., Liu, Y.J.: A review of blockchain technology. *Comput. Eng.* **45**(05), 1–12 (2019)
9. Buterin, V.: A next-generation smart contract and decentralized application platform. *White Pap.* **3**(37) (2014)
10. Cao, N.: Blockchain technology in the era of rule of law to help financial precision poverty alleviation. *J. Soc. Sci. Harbin Normal Univ.* **9**(05), 87–89 (2018)



An Intent-Based Routing Scheme in Satellite IoT

Guanghai Zhang¹, Kang Liu¹, Xue Chen^{2,3}(✉), Xu Liu¹, Zhen Pan¹,
and Tianyang Jiang²

¹ Power Dispatching Control Center of Guizhou Power Grid Co., Ltd., Guiyang, China

² University of Electronic Science and Technology of China, Chengdu, China
chenxue61@163.com

³ Tianfu Collaborative Innovation Center, University of Electronic Science and Technology of
China, Chengdu, China

Abstract. As the satellite IoT integrates the satellite network and the IoT, a large number of IoT services are transmitted through satellite links, which is easy to cause the congestion of satellite links during packet routing. By incorporating the concept of intent, this paper proposes an intent-based routing scheme in satellite IoT to solve the congestion problem of satellite link. The routing scheme consists of two parts: the packet preprocessing and shortest path search. In the stage of packet preprocessing, we construct an intent-based system model and design the *packet preprocessing subalgorithm* based on intent. It is used to analyze the user demand, data priority and network congestion state through the intent. In the stage of shortest path search, we design the *shortest path search subalgorithm* based on the interstellar link and user intent. When packet is routed between satellites, the interstellar links and logical locations between the source and destination satellites jointly determine the next hop of satellite. The simulation results prove that the proposed routing scheme has good performance in average end-to-end delay, packet loss rate and network throughput, which can greatly reduce the congestion of satellite links.

Keywords: Satellite IoT · Intent · Routing · Congestion

1 Introduction

In recent years, the IoT has become an indispensable part of people's daily life [1]. However, the IoT is a kind of terrestrial network. It is greatly affected by terrain conditions. And difficult to cover the remote and untraversed areas, such as the mountains and seas. Especially in the 5G network or 6G in the future, people expect to be able to get whatever service they want regardless of their location (whether in the city center or remote rural areas) [2]. Fortunately, the satellite communications can solve the any time and any place coverage problem to some degree [3, 4]. Thus, the concept of satellite IoT came into being, which is integrated by the satellite network and the IoT [5, 6].

Figure 1 shows the traditional satellite IoT [7]. Satellites are configured on the ground when they fly over a site. If the configurations need to be changed or cancelled, it can only

be done manually. However, these manual operations are not only inefficient, but also easily cause network congestion during packet routing. And in the traditional satellite IoT, the IoT services must be transmitted through satellite links. Due to the variety of IoT services, the data volume is large and complex. This easily causes satellite link congestion during packet routing. Therefore, real-time detection of satellite link status is required during packet routing. In this way, the best data packet transmission path can be found to avoid satellite link congestion.

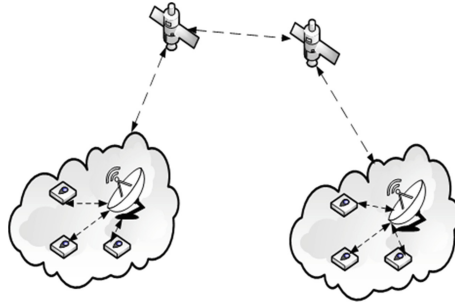


Fig. 1. The traditional satellite IoT.

In this paper, aiming at the congestion problem in satellite IoT, we propose the Intent-Based Routing algorithm (IBR) using the Intent Based Network (IBN) technology. The IBR algorithm consists of two parts: the packet preprocessing and shortest path search. The main contributions of this paper are as follows.

- To preprocess data, an intent-based satellite IoT system model (ISI) is proposed. And we design the packet preprocessing subalgorithm based on intent in IBR. When the packet is received by satellites, the current network state information is detected in real time through the ISI model, and the link state of the packet transmission is analyzed. When a link is congested, IBR performs rollback operations based on the priority of the resolved service type, to reduce the packet loss rate of important packets and improve the network throughput.
- In IBR, the shortest path search subalgorithm based on the interstellar link is designed. After the data packet is preprocessed, a route with minimum transmission delay is found based on the interstellar link model. Then, the logical location of the source/destination satellites and the interplanetary link together determine the next hop of the packet route. It achieves a smaller transmission distance between the source and destination location of the packet and reduces the end-to-end delay of the packet.

2 Related Work

In recent years, with the application of various advanced communication technologies in many fields such as data center network [8], wireless access network [9] and optical

network [10], satellite IoT has been greatly developed. For example, in [11], the authors proposed a software defined satellite network. It solved the problem of dynamic change of satellite network topology and limited on-board processing capacity; In [12], the authors proposed a OpenSAN satellite network, which separated the data plane (satellite infrastructure) from the control plane (geosynchronous orbit group). It provided a good idea for the design of multi-layer space network; In [10], the authors discussed the payload of satellite and solved the problems existing in the current construction of space information system.

However, the above studies focused on the space network without considering the idea of the integrated automation of heaven and earth. In order to further improve the problem of network automation, the concept of IBN had been introduced [13] in satellite IoT. In 2015, the ONF organization proposed the concept of “network intent” [14]. And in 2017, Gartner formally defined the IBN [15]. In IBN, intelligent software determines how to translate intents into infrastructure-specific policies that enable the network to operate in the desired manner. However, the traditional satellite IoT requires administrators to enter specific commands to configure manually. In the intent-based network, the administrator can directly input the state to be performed automatically by the network, rather than by the administrator entering program commands step by step. Thus, the network automation of satellite IoT is expected to reduce satellite link congestion by using IBN technology.

In addition, a good routing algorithm plays a decisive role in alleviating link congestion. Nowadays, some existing routing algorithms of satellite IoT have developed into a connection oriented network idea. Its specific idea is that the network control center uniformly calculates the satellite routing. Thus, the satellite routing table can be formulated according to the network link, and the data packets are transmitted according to the routing table. However, the satellite is moving at a high speed, and the network topology changes at any time. Thus, the pre-calculated routing path may not be optimal as satellite nodes move. In order to solve this problem, many scholars have conducted in-depth research on routing algorithm. For example, in [16], the authors proposed the location-based routing (LBR) algorithm to reduce the routing overhead. Although the LBR algorithm generated the minimum propagation delay path, it ignored the existence of reverse rotation slot; In [17], a low complexity routing algorithm based on load balancing was proposed. It obtained the satellite congestion state through mutual communication between satellites. However, it increased the processing burden of satellites and did not consider the priority of users; In [18], a software defined routing algorithm was proposed. And the optimal path could be obtained by collecting the satellite congestion state in real time. However, the QoS and traffic classification were not well guaranteed; In [19], the authors proposed another software defined routing algorithm (SDRA). It could obtain the optimal routing path by centralized routing strategy. However, the SDRA algorithm was based on the fixed network topology. It putted forward higher requirements for the storage capacity of satellite.

Thus, we need to design a more efficient satellite routing algorithm, which can realize automatic configuration through IBN and solve the problem of satellite link congestion at the same time.

3 Model Building

In this section, we describe the ISI system model and interstellar link model, which are used to guide the design of IBR algorithm.

3.1 ISI System Model

In the stage of packet preprocessing in IBR algorithm, the ISI system model is first built based on IBN technology. The ISI model is shown in Fig. 2. According to the function of Satellite IoT, we divide the system model into three planes: the application layer, the control layer, and the data forwarding layer.

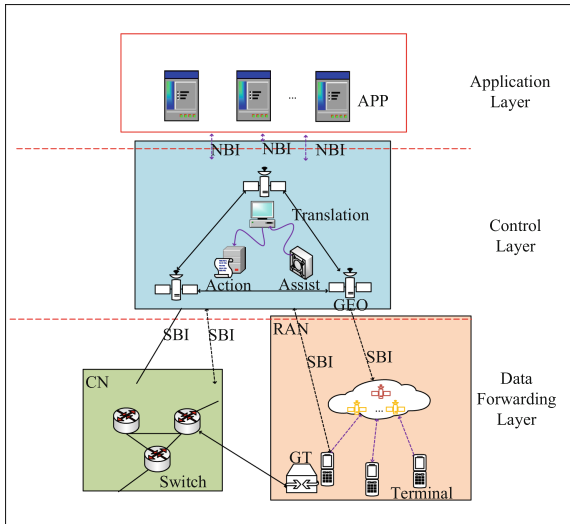


Fig. 2. The ISI system model.

In ISI model, it not only cares about the upper-layer services and applications, but also pays attention to how the underlying network is set up and run. This greatly increases the inflexibility and controllability of the network. The purpose of ISI model is to convert user intents (or needs) into actual strategies. That is, the intent drives the planning and design of the network, thereby saving network resources. Thus, the intent controller is the key of ISI model. It realizes the real-time detection of the network status. The intent controller is deployed in control layer. It allows all types of applications to express their needs and constraints in their own terms. Then it converts the application requirements to the underlying network requirements. Thus, when intent controller performs a task, it can understand the needs of the application and provide the required services through the management of the underlying network resources.

(1) Application layer

The application layer of ISI consists of several applications, such as routing, bandwidth allocation. These applications can submit user intents to the intent controller via the northbound interface (NBI). The NBI is between in the application layer and control layer. By using NBI, we can program the network device, and receive feedback from the control layer or the abstract model of underlying network. Therefore, the main function of application layer is to provide the requirements so that users can be closer to the control network.

(2) Control layer

The control layer of ISI model is in charge of collecting the network status information. It is mainly composed of intent controllers. Due to the large coverage of geosynchronous orbit satellite (GEO) [20], three GEOs can cover the entire earth and collect entire network status information. Thus, we employ the intent controller on the GEO. After the intent controller receives the user intents from the application layer, it translates these intents into implementable policies. The intent controller consists of three main modules: translation, action, and assist, as shown in Fig. 3. After the intent is captured, intent controller forms a closed loop to ensure that the intent is not disturbed by the sudden situation.

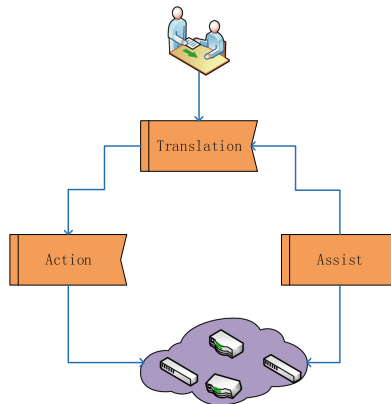


Fig. 3. The intent controller.

(3) Data layer

The data layer separates the control decision from the forwarding hardware and forwards the data under the control of the intent controllers. In data layer, as the medium/low orbit satellite has a lower latency and cost compared to the GEO, we use the medium/low orbit satellites as forwarding planes. However, the medium/low orbit satellites in data layer only have the capability of data forwarding. The data is forwarded to the underlying devices via the interstellar links in medium/low orbit satellites. At the same time, when a user wants to communicate with other users or devices with a long distance, the

user can upload the data from the low orbit satellite, and then forward the data using the medium/low orbit satellite links.

3.2 Interstellar Link Model

The interstellar link model is shown in Fig. 4. It consists of N orbital planes. Each orbital plane includes K satellites. Each satellite has a logical address S_{ij} , which denotes its routing address. Here, the i represents the i -th orbital plane and i is not more than N (that is, $0 < i \leq N$). The j represents the j -th satellite in the i -th orbital plane and j is not more than K (that is, $0 < j \leq K$). There exist two kinds of interstellar links, including intra-plane link L_{intra} , and inter-plane link L_{inter} . In Fig. 4, each satellite theoretically involves four interstellar links: two L_{intra} links and two L_{inter} links. However, as the satellite of polar region operates a high speed, the antenna system cannot track the satellite's position in real time. This leads that inter-plane links need to be disconnected and reconnected in polar region. Therefore, the satellites of polar region have two L_{intra} links. In addition, there exists a counter-rotating seam in satellite network (The red dotted line in Fig. 4). It is between in two orbital planes that are moving in opposite directions. The inter-plane between the two planes immediately adjacent to the counter-rotating seam cannot be used for link acquisition. Thus, there are only three satellite links in these two planes, including two L_{intra} links and one L_{inter} link.

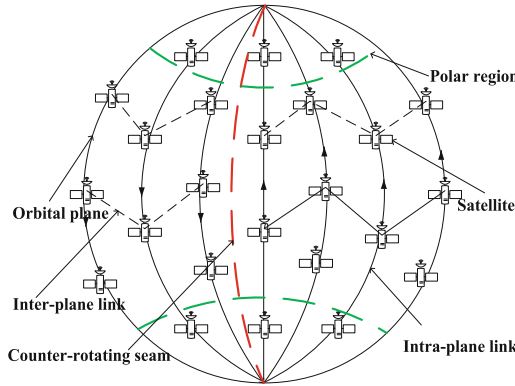


Fig. 4. Interstellar link model. (Color figure online)

From the Fig. 4, we can get that the length of L_{intra} is fixed. It can be calculated according to Eq. (1).

$$L_{intra} = \sqrt{2}(R + h) \sqrt{1 - \cos\left(\frac{360}{K}\right)} \quad (1)$$

where R is the radius of the earth, and h is the height of the orbital plane from the earth.

The inter-plane link L_{inter} can be calculated by Eq. (2).

$$L_{inter} = \sqrt{2}(R + h) \sqrt{1 - \cos\left(\frac{360}{2 \times N}\right) \times \cos(D)} \quad (2)$$

where D is the latitude of the inter-plane. And the main symbols adopted in this work are illustrated in Table 1 for reference.

Table 1. Summary of key notations

Notation	Description
N	Number of orbital planes
K	Number of satellites in each orbital plane
S_{ij}	The i -th orbital plane and j -th satellite's logical address
L_{intra}	The length of the intra-plane link in interstellar
L_{inter}	The length of the inter-plane link in interstellar
R	The radius of the earth
h	The height of the orbital plane from the earth
D	The latitude of the inter-plane
C_k	The capacity of the k -th satellite link
L_{avg}	The link average load
T_h	The link average load threshold
L_s	The latitude values of the source satellite
L_d	The latitude values of the destination satellite

4 Routing Scheme Design

Due to the large number of services in satellite IoT, when the services are transmitted through the satellite links, it easily causes the problem of satellite link congestion. Thus, the network topology needs to be checked for congestion before routing. Therefore, the IBR algorithm is designed based on this idea. It can detect various status information of the network in real time and view the link congestion status when routing. The IBR algorithm is divided into two stages: the packet preprocessing and shortest path search. In the stage of the packet preprocessing, the ISI model first detects the network status, and obtains the user's type and requirements by parsing intent. Then it determines the actual operations. In the stage of shortest path search, a path with the shortest transmission distance is determined according to the interstellar link and logical location between the source satellite and destination satellite.

Because link resources and capacity are limited, it is necessary to evaluate the link average load at control layer before finding the shortest path. Suppose the capacity of a satellite link is C_k . Thus, we get that the capacity array of n satellite links is $C = \{C_1, C_2 \dots C_{n-1}, C_n\}$. The link average load L_{avg} is written as Eq. (3).

$$L_{avg} = \frac{\sum_1^n C_k}{n} \quad (3)$$

We define a link average load threshold T_h , as shown the Eq. (4):

$$T_h = \min\{C_k\}, k \in [1, n] \quad (4)$$

L_{avg} and T_h are two parameters of IBR algorithm. When the data packet comes, the IBR algorithm calculates L_{avg} , and compares L_{avg} with T_h to perform different routing strategies.

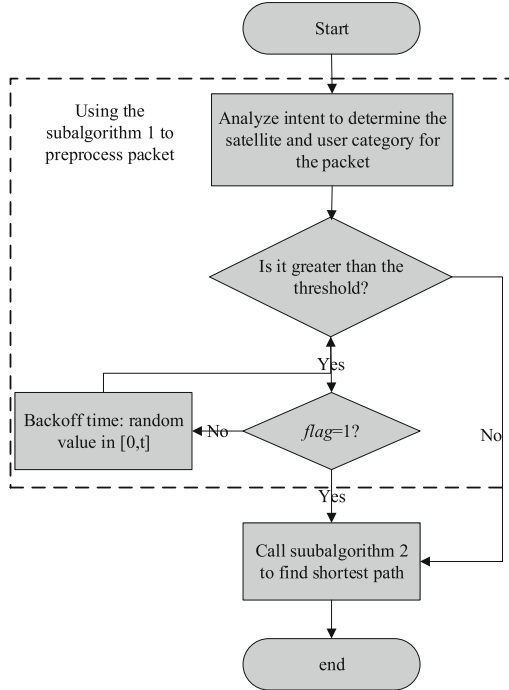


Fig. 5. The flowchart of IBR algorithm.

Figure 5 shows the flow chart for IBR algorithm. To guarantee user's QoS, it preprocesses the upcoming data packets by calling the *subalgorithm 1*. When packet is accessed, the intent control layer calculates the link average load L_{avg} and packet type. In this paper, we divide the data packets into two types using value $flag = \{1, 0\}$. $flag = 1$ represents high priority data packet and $flag = 0$ means low priority data packet. Each data packet carries an ID, which is the identity for determining the priority of packet. Then we compare the link average load L_{avg} with the threshold value T_h . When L_{avg} is less than T_h , the two types of data packets directly perform operation of shortest routing search. When L_{avg} is greater than or equal to T_h , the high priority data packet performs the operation of shortest routing search, and the low priority data packet performs the back-off. The back-off time is a random value in $[0, t]$. After waiting the back-off time, the control layer decides whether the operation of shortest routing search is performed by calling the *subalgorithm 2*. *Subalgorithm 2* is to determine the moving direction and next hop of the data packet, and finally reach the destination address.

Subalgorithm 1 describes the operation of packet preprocessing in detail. It includes four main steps: satellite logical area division, intent analysis, network monitoring and route formulation. First, we logically divide the satellite's regional location to determine the control area of each satellite. When a data packet is transmitted through the satellite IoT, the corresponding intent of submission request is determined. Secondly, it analyzes the intent to obtain the information $\{flag, T_h\}$ related to the data packet. Thus, we can determine the specific content and priority of the data packet, as well as the link congestion threshold. Then the ISI model detects the congestion status of the network and obtains the link average load L_{avg} . When $L_{avg} \geq T_h$, it indicates that satellite link congestion has occurred. In this case, data packets may not be transmitted successfully. And it will significantly increase the congestion and packet loss rate. Therefore, in *subalgorithm 1*, we transmit high-priority data packets directly. For the low-priority data packets, we perform the random time backoff. It means to wait for a period of time ($0-t$ s) before transmitting. When $L_{avg} < T_h$, it means that the satellite link is relatively idle and the number of transmitted data packets is relatively small. Thus, data packets can be directly transmitted after being queued according to priority. By preprocessing the upcoming data packets, the IBR algorithm can effectively reduce the packet loss rate and improve the throughput.

Subalgorithm 1: Packet Preprocessing Algorithm

```

1: Begin
2: Receive a packet;
3: Get  $L_{avg}$ ,  $flag$  and  $T_h$  of the packet;
4: if ( $L_{avg} \geq T_h$ )
5:   if ( $flag == 1$ )
6:     Send packet to the next satellite directly;
7:   else if ( $flag == 0$ )
8:     Wait for  $0-t$  seconds randomly and send the packet;
9:   end if
10: else if ( $L_{avg} < T_h$ )
11:   Packets are queued for priority;
12: end if
13: end

```

For the *subalgorithm 2*, there exists three cases to get the shortest path using the interstellar link model.

Case 1: Both the source satellite and the destination satellite are in the polar region. It means $L_S > 70^\circ$ & $L_d > 70^\circ$. Here, L_S and L_d are latitude values of the source and destination satellite respectively.

Case 2: The source satellites and destination satellites are not in the polar region. It means $L_S < 70^\circ$ & $L_d < 70^\circ$.

Case 3: One of the source/destination satellites is in the polar region and the other is outside the polar region.

For case 1, we need to determine whether the source and destination satellites are in the same satellite orbital plane. Because only intra-plane link L_{intra} exists in the polar region. There are no inter-plane link L_{inter} . If the source satellite and destination satellite are in the same orbital plane ($L_s = L_d$), the data packet can reach the destination satellite from the source satellite along the link L_{intra} in the plane by the path $Ps = \{S_{i_s, j_s}, S_{i_s, j_s+1}, S_{i_s, j_s-1}, S_{i_s, j_d}\}$. If the source satellite and destination satellite are not in the same orbital plane, the data packet first needs to move along the link L_{intra} to near-polar earth plane. Then the packet is transmitted to the satellite in the same orbital plane as the destination satellite. Thus, the packet can finally reach to the destination satellite by the path $Ps = \{S_{i_s, j_s}, S_{i_s, 1}, S_{i_d, 1}, S_{i_d, j_d}\}$, thus completing the entire transmission request.

For case 2, as the source satellite and destination satellite are not in the polar region, both link L_{intra} and L_{inter} can be used. As the higher the latitude of the satellite, the shorter the interstellar link between planes. Thus, the packet first moves vertically to the satellite in the same orbital plane as the destination satellite, and then moves vertically toward the destination satellite as the shortest path. When the latitude of source satellite is higher than destination satellite ($L_s \geq L_d$), the data packet is first transmitted along the shorter link L_{inter} , and then along the link L_{intra} towards to the destination satellite. Thus, the path Ps is $\{S_{i_s, j_s}, S_{i_d, j_s}, S_{i_d, j_d}\}$. When $L_s < L_d$, the data packet is first transmitted along the link L_{intra} in the plane, and then transmitted along the link L_{inter} between the planes, so as to ensure the shortest path length. Thus, the Ps is $\{S_{i_s, j_s}, S_{i_s, j_d}, S_{i_d, j_d}\}$.

For case 3, whether the destination satellite or source satellite is within the polar region, the data packets must reach the near-polar plane along the link L_{intra} . After the data packet reaches the near-polar satellite, we check whether it is in the same orbit as the destination satellite. If $L_s \geq L_d$, the data packet can be transmitted directly to the destination satellite along the link L_{intra} by the path $Ps = \{S_{i_s, j_s}, S_{i_d, j_s}, S_{i_d, j_d}\}$; Otherwise, the data packet is routed along the link L_{inter} in the near-polar plane, reaching the satellite in the same orbital plane as the destination satellite, and finally according to the link L_{intra} for transmission. Thus, the Ps is $\{S_{i_s, j_s}, S_{i_s, j_s+1}, S_{i_s, j_s-1}, S_{i_d, j_d}\}$.

Based on the above description, for different packet types and different locations of source/destination satellite, the IBR algorithm will perform different paths strategies. At the same time, the ISI model checks the congestion of the satellite link before routing. This way can greatly avoid the congestion caused by the influx of data packets into satellite links. Therefore, the IBR algorithm has the advantages of high throughput, low delay, and high packet loss rate.

Subalgorithm 2: Shortest Path Search Algorithm**Input:** $S_{i_s, j_s}, S_{i_d, j_d}, L_s, L_d$, set $P_s = \emptyset$ as initial path**Output:** the shortest path P_s

```

1: if ( $L_s > 70^\circ$  &&  $L_d > 70^\circ$ ) // (case 1)
2:   if ( $L_s == L_d$ )
3:      $S_{i_s, j_s} \rightarrow S_{i_s, j_s \pm 1} \rightarrow S_{i_s, j_d}$ ;
        $P_s = \{ S_{i_s, j_s}, S_{i_s, j_s + 1}, S_{i_s, j_s - 1}, S_{i_s, j_d} \}$ ;
4:   else
5:      $S_{i_s, j_s} \rightarrow S_{i_s, 1} \rightarrow S_{i_d, 1} \rightarrow S_{i_d, j_d}$ ;
        $P_s = \{ S_{i_s, j_s}, S_{i_s, 1}, S_{i_d, 1}, S_{i_d, j_d} \}$ ;
6:   end if
7: else if ( $L_s < 70^\circ$  &&  $L_d < 70^\circ$ ) // (case 2)
8:   if ( $L_s \geq L_d$ )
9:      $S_{i_s, j_s} \rightarrow S_{i_d, j_s} \rightarrow S_{i_d, j_d}$ ;
        $P_s = \{ S_{i_s, j_s}, S_{i_d, j_s}, S_{i_d, j_d} \}$ ;
10:  else
11:     $S_{i_s, j_s} \rightarrow S_{i_s, j_d} \rightarrow S_{i_d, j_d}$ ;
        $P_s = \{ S_{i_s, j_s}, S_{i_s, j_d}, S_{i_d, j_d} \}$ ;
12:  end if
13: end if
14: else // (case 3)
15:   if ( $L_s == L_d$ )
16:      $S_{i_s, j_s} \rightarrow S_{i_s, j_s - 1} \rightarrow S_{i_s, j_d}$ ;
        $P_s = \{ S_{i_s, j_s}, S_{i_s, j_s - 1}, S_{i_d, j_d} \}$ ;
17:   else
18:      $P_s = \{ S_{i_s, j_s}, S_{i_s, j_s + 1}, S_{i_s, j_s - 1}, S_{i_d, j_d} \}$ ;
19:   end if
20: end if

```

5 Simulation and Results

In this section, by OMNeT++ simulation, we compare the performance of the IBR algorithm with the existing algorithms DRA [21] and LCPR [22]. In our simulation, we don't consider the satellites of polar region. Thus, each satellite in our simulation is assigned for two L_{intra} links and one L_{inter} link. The track height and inclination of satellite are set at 780 km and 86.4° , respectively. The capacity of ISL is 20 Mbps. In ISI, the data transmission rate is set to 1–2 Mbps. Each packet size is 1024 bytes. And the data packet obeys the Poisson distribution and is generated with the speed of 300 packets per second. Both high priority packets and low priority packets are subject to Poisson distribution.

The Fig. 6 has displayed the relationship between packet loss rate and data transmission rate for IBR, DRA and LCPR algorithms. The average packet loss rate of IBR is about 4% lower than LCPR algorithm, and 10% lower than that of DRA algorithm. When the data transmission rate is lower than 1.4 Mbps, the packet loss rate of IBR is lower than 3.5% and the growth rate is slow. This is because we queue the data packets and introduce random back-off time to effectively reduce the packet loss rate. When the data transmission rate exceeds 1.5 Mbps, more data packets enter the satellite link. Thus, the impact of link congestion on the packet loss rate increases gradually, resulting in a positive correlation between the packet loss rate and the data transmission rate.

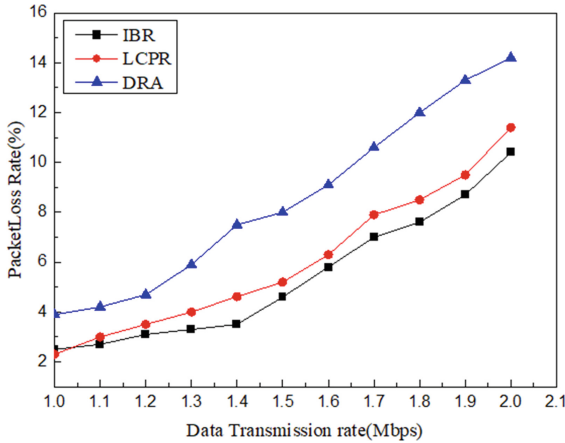


Fig. 6. The relationship between packet loss rate and data transmission rate.

Figure 7 shows the performance of end-to-end delay. IBR algorithm optimizes the average end-to-end delay by 7% and 11% respectively compared with LCPR algorithm and DRA algorithm. Because ISI can monitor the network status and adjust the routing path in real time. When the link is idle, data packets will be queued for transmission.

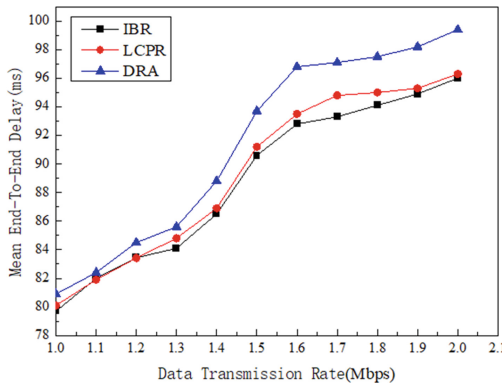


Fig. 7. The relationship between average end-to-end delay and data transmission rate.

When the link congestion occurs, high priority packets have no waiting time and can be transmitted first, which effectively improves the average end-to-end delay.

In Fig. 8, our experiments compare throughput performance. The throughput is also positively correlated with the data transmission rate. When the data transmission rate increases, the data packet increases and the throughput also improves. IBR algorithm increases the average throughput by 5.8 Kbps and 11.6 Kbps compared with LCPR and DRA algorithms, respectively. Thus, the IBR algorithm has a significant improvement in terms of throughput.

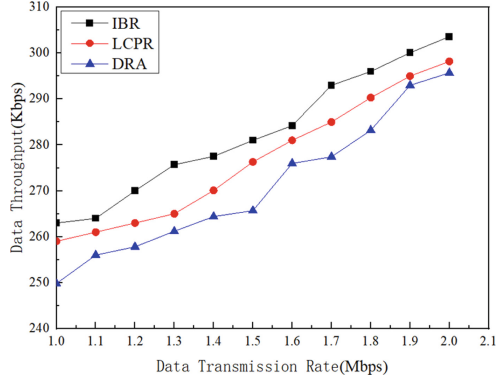


Fig. 8. The relationship between throughput and data transmission rate.

6 Conclusion

Aiming at the problem of satellite link congestion caused by the large number of service types in satellite IoT, we have proposed the ISI system model and designed an effective routing algorithm TBR by using INB technology. The IBR routing algorithm is mainly designed based on the ISI system model. The ISI model realizes the automatic configuration and security monitoring of the underlying network by parsing the business intent. In TBR algorithm, after network status detection through ISI model, the operation of data packet access or waiting is decided according to the congestion state of satellite link. Then, a shortest path is found according to the logical position of satellites and the interstellar link between satellites. However, the proposed routing algorithm only considers the single-layer satellite network, but the actual satellite IoT is complicated. Therefore, we will consider more complex satellite routing models, and design flexible and universal routing strategies in the future.

Acknowledgement. This research was partially supported by National Natural Science Foundation of China (61971105), Key Science and Technology Project of China Southern Power Grid Co., Ltd. (GZKJXM20210011), Sichuan science and technology program (2021YFG0150, 2021YFG0037, 2022YFG0173).

References

1. Kodheli, O., Maturo, N., Shankar, B., et al.: Satellite communications in the new space era: a survey and future challenges. *IEEE Commun. Surv. Tutor.* **23**, 70–109 (2021)
2. Zhu, X., Jiang, C.: Integrated satellite-terrestrial networks toward 6G: architectures, applications, and challenges. *IEEE Internet of Things J.* **9**, 437–461 (2022)
3. Yan, Y., Guang, H., Hui, X.: A survey on secure routing protocols for satellite network. *J. Netw. Comput. Appl.* **145**, 1–17 (2019)
4. 2017 SIA State of Satellite Industry Report. <http://www.sia.org/annual-state-of-the-satellite-industry-reports/2017-sia-state-of-satellite-industry-report/>. Accessed on 2017
5. Hassan, N.U.L., Huang, C., Yuen, C., Ahmad, A., Zhang, Y.: Dense small satellite networks for modern terrestrial communication systems: benefits, infrastructure, and technologies. *IEEE Wirel. Commun.* **27**, 96–103 (2020)
6. Gao, Y., Ao, H., Zhou, Q., et al.: Modeling of satellite communication systems design with physical layer security. In: *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 1680–1683 (2018)
7. Lu, Y., Min, G., Zuo, Z., Liang, R., Duan, Z.: Structural performance of satellite networks: a complex network perspective. *IEEE Syst. J.* **15**, 3848–3859 (2021)
8. Jain, S., Kumar, A., Mandal, S., et al.: B4: experience with a globally-deployed software defined WAN. In: *ACM SIGCOMM Conference*, Hongkong, pp.3–14 (2013)
9. Patel, A.N., Ji, P.N., Wang, T.: QoS-aware optical burst switching in OpenFlow based software-defined optical networks. In: *17th International Conference on Optical Networking Design and Modeling (ONDM)*, Brest, pp. 275–280 (2013)
10. Yang, X.N., Xu, J.L., Lou, C.Y.: Software-defined satellite: a new concept for space information system. In: *2nd International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, Harbin, pp. 586–589 (2012)
11. Gudipati, A., Perry, D., Li, L.E., et al.: SoftRAN: software defined radio access network. In: *2nd ACM SIGCOMM Workshop Hot Topics in Software Defined Networking*, Hongkong, pp. 25–30 (2013)
12. Bao, J.Z., Zhao, B.K., Yu, W.R.: OpenSAN: a software-defined satellite network architecture. In: *SIGCOMM Conference*, Chicago, pp. 347–348 (2014)
13. Intent-Based Networking-Building the bridge between business and IT. <https://www.cisco.com/c/en/us/solutions/intent-based-networking.html>. Accessed on 2017
14. Intent NBI. <https://www.opennet-working.org/>. Accessed on 2016
15. Mohammed, A., Gharbaoui, M., Martini, B., et al.: SDN controller for network-aware adaptive orchestration in dynamic service chaining. In: *IEEE NetSoft Conference and Workshops*, pp. 126–130 (2016)
16. Ding, J.D., Zhang, Y., Li, R.N., et al.: A distributed routing algorithm for LEO satellite networks. In: *International Conference on Human Centered Computing*, Kazan, pp. 258–264 (2017)
17. Liu, X.M., Yan, X.M., Jiang, Z.Q., et al.: A low-complexity routing algorithm based on load balancing for LEO satellite networks. In: *Vehicular Technology Conference*, Boston, pp. 1–5 (2015)
18. Zhang, L.L., Wang, X.W., Huang, M.: A routing scheme for software-defined satellite network. In: *2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, Guangzhou, pp. 24–31 (2017)
19. Guo, A.D., Zhao, C.L., Xu, F.M., et al.: LEO satellite routing algorithm in software defined space terrestrial integrated network. In: *17th International Symposium on Communications and Information Technologies (ISCIT)*, Cairns, pp. 1–6 (2017)

20. Cai, Y., Zhang, S., Hou, J., Jiao, H.: Internet of Things over GEO satellite: a novel space information network solution. In: 2021 International Wireless Communications and Mobile Computing (IWCMC), pp. 177–181 (2021)
21. Ekici, E., Akyildiz, I., Bender, M.: Datagram routing algorithm for LEO satellite networks. In: Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 500–508 (2000)
22. Liu, H.Y., Sun, F.C., Yang, Z.A., et al.: A novel distributed routing algorithm for LEO satellite network. In: International Conference on Industrial Control and Electronics Engineering (ICICEE), Xi'an, pp. 37–40 (2012)

Game Theory



Interference Management in Terrestrial-Satellite Networks Using Stackelberg Game

Yaomin Zhang^{1,2}, Haijun Zhang^{1,2(✉)}, and Keping Long^{1,2}

¹ University of Science and Technology Beijing, Beijing, China
yaominzhang@xs.ustb.edu.cn, haijunzhang@ieee.org, longkeping@ustb.edu.cn

² Beijing Engineering and Technology Research Center for Convergence Networks
and Ubiquitous Services, Beijing, China

Abstract. The terrestrial-satellite network (TSN) has been recognized a potential network realizing wide-seamless coverage and high transmission rate. In this paper, a Stackelberg game model for the interference management problem is studied in spectrum sharing TSNs, where the competition between the leaders and followers is recognized as the utility maximization between the satellite user (SU) and terrestrial users (TUs). In the leader sub-game, the utility function of SU is constructed of transmission rate and cross-tier interference reacting to the followers, and the optimal strategy is acquired by mathematical derivation. Taking account to the fairness among TUs, the maximization of the minimum worst-case transmission rate is established in the follower sub-game by optimizing the cooperative scheduling and power allocation strategy jointly. Then we propose an evolutionary gale-shapley algorithm to solve the cooperative scheduling problem. And the closed-form expression of TU power is derived by the Lagrangian dual decomposition approach. By the Stackelberg iteration of the leader strategy and follower strategy, the Stackelberg equilibrium point is ultimately obtained. We give the performance simulation to verify the convergence and effectiveness of the proposed algorithm.

Keywords: Terrestrial-satellite network · Interference management · Stackelberg game · User fairness · Cooperative scheduling

This work is supported in part by National Key R&D Program of China (Grant No. 2020YFB1806103), in part by Beijing Natural Science Foundation (L212004), in part by the State Key Laboratory of Advanced Metallurgy under Grant KF20-04, in part by the Fundamental Research Funds for the Central Universities under Grants FRFTP-19-002C1 and RC1631, and in part by Beijing Top Discipline for Artificial Intelligent Science and Engineering, University of Science and Technology Beijing.

1 Introduction

The terrestrial-satellite network (TSN) plays a prominent role in the 5G networks and beyond for supporting the ubiquitous wireless services with wide coverage and large-capacity [1–3]. In the TSN, the users can choose the satellite or the traditional terrestrial base stations (BSs) based on their demand, which improves the network capacity and flexibility. However, co-existing satellite, terrestrial BS, and users make the wireless resource management more complicated [4]. And the shortage of spectrum resources has also brought pressure to performance improvement in the TSN [5]. The researchers have done a lot work to deal with these challenges. In [6], the authors considered the spectrum sharing between the satellite-tier and terrestrial-tier, and proposed an energy efficient power allocation approach to satisfy the diverse delay quality-of-service (QoS) requirements. The paper [7] investigated the spectrum access problem in terrestrial-satellite heterogeneous networks by nonorthogonal multiple access (NOMA) for each beam, where satellite users (SUs) and terrestrial users (TUs) used the same spectrum in the underlay mode to improve spectrum utilization. The user scheduling and resource allocation problem was studied for the spectrum coexisting based TSN in [8], and the authors proposed a joint optimization scheme to improve system sum rate and energy efficiency.

Although spectrum sharing and NOMA can increase spectral efficiency of TSN, the serious co-channel interference is inevitable [9], especially in the multiple BSs based TSNs. Resource allocation and interference management has been considered to coordinate the user access and allocation of resource blocks, so as to achieve desirable performance [10–13]. In [10], the authors studied interference efficiency maximization problem in a two-tier heterogeneous network and proposed a robust resource allocation scheme to achieve the rate-interference tradeoff. In [11], the authors proposed to mitigate the cross-tier interference in the cloud based TSN by using the channel information at the cloud, a two-tier Stackelberg game between cloud and users was then established to obtain the optimal pricing and resource allocation strategy. The minimization problem of the beam usage and power consumption while ensuring the user demand was investigated in [12], where the optimized user interference was within a tolerable range. Also in [13], a coordinated multi-point (CoMP) transmission and NOMA based resource allocation model was designed to mitigate the inter-cell interference, in which the authors discussed the different CoMP strategy and derived the optimal power solutions among the coordinating BSs to improve the spectral efficiency.

However, the cross-tier interference problem by jointly considering the QoS guarantee, user fairness, cooperative scheduling, and inter-cell interference mitigation has not been studied systematically in multiple BSs based two-tier TSN. In this paper, the joint resource management problem for utility maximization is investigated in the terrestrial-satellite two-tier networks. With the consideration of the QoS, cross-tier interference, and the user fairness, we model the optimization problems of satellite-tier and terrestrial-tier, respectively. The two-tier joint optimization problem is then formulated as a Stackelberg game between

the SUs and TUs, where SUs are the leaders and TUs are the followers. In the satellite-tier optimization problem, the leaders choose the optimal power strategy to guarantee their high transmission rate without serious cross-tier interference to the followers. Based on the strategy of the leaders, the terrestrial-tier maximizes the minimum worst-case transmission rate among the followers to achieve the fairness of the TUs. The optimization problem of follower sub-game includes two nested problems: cooperative scheduling problem and power allocation problem. An evolutionary gale-shapley algorithm is proposed to solve the cooperative scheduling problem. The power allocation problem is then transformed into a convex optimization problem, and we derive the closed-form expression of the optimal power strategy of followers. By the Stackelberg iteration of the leader strategy and follower strategy, the Stackelberg equilibrium point is ultimately obtained.

2 System Model and Stackelberg Game Formulation

2.1 System Model

The considered downlink TSNs consist of a LEO, K terrestrial BS with $\mathbb{K} = \{1, 2, \dots, K\}$, U^T TUs with $\mathbb{U}^T = \{1, 2, \dots, U^T\}$, and U^S SUs with $\mathbb{U}^S = \{1, 2, \dots, U^S\}$. The system bandwidth is B , which is divided to N orthogonal subchannels (SC) and each SC is identical for a SU and a terrestrial cell for spectrum sharing. To guarantee the QoS of satellite-tier, each SU communicates with LEO in different SC. And the NOMA is adopted in the terrestrial-tier to improve the spectrum utilization.

Thus, in a terrestrial-satellite system, the signal received of the SU u is

$$y_u^S = \underbrace{g_u^S \sqrt{p_u^S} s_u^S}_{\text{Desired Signal}} + \underbrace{\sum_{k \in \mathbb{K}} \left(g_{k,u}^T \sum_{j \in \mathbb{K}^T} x_{k,j}^T \sqrt{p_{k,j}^T} s_{k,j}^T \right)}_{\text{Cross-tier Interference}} + n_u^S, \quad (1)$$

where g_u^S , $g_{k,u}^T$, p_u^S , and $p_{k,j}^T$ are the channel coefficient from the LEO to SU u , the channel coefficient from the terrestrial BS k to SU u , the transmission power from the LEO to SU u , and the transmission power from the terrestrial BS k to SU j , respectively. s_u^S and $s_{k,j}^T$ denote the transmission signal from the LEO to the SU u , and the transmission signal from the BS k to the TU j . n_u^S is the additive white Gaussian noise (AWGN) of the satellite-tier. Denote that $E[|s_u^S|^2] = 1$, $E[|s_{k,j}^T|^2] = 1$, and $E[n_u^S] = \sigma^S$. We define the cooperative factor $x_{k,u}^T$, where $x_{k,u}^T = 1$ indicates the TU u is served by BS k , $x_{k,u}^T = 0$, otherwise.

Assume that BS k and SU u communicate with the same SC, then the signal to interference noise ratio (SINR) of the SU u is

$$\gamma_u^S = \frac{|g_u^S|^2 p_u^S}{|g_{k,u}^T|^2 \sum_{j \in \mathbb{M}_k} x_{k,j}^T p_{k,j}^T + \sigma^S}, \quad (2)$$

where $\mathbb{M}_k = \{x_{k,u}^T = 1 | u \in \mathbb{U}^T\}$ represents the set of users served by BS k . Let $B_{sc} = B/N$ denote the bandwidth of each SC, N_0^S denote the noise power spectral density of the satellite-tier, and we have $\sigma^{S^2} = B_{sc}N_0^S$. Based on the Shannon theorem, the transmission rate of the SU u is

$$C_u^S = \log_2(1 + \gamma_u^S). \tag{3}$$

For terrestrial system, the users served by the same BS form a NOMA cluster, thus the intra-cluster interference caused by NOMA mechanism should be considered. The received signal of the TU u includes desired signal, intra-cluster interference, cross-tier interference, and noise, which is

$$\begin{aligned}
 y_u^T = & \underbrace{\sum_{k \in \mathbb{K}} x_{k,u}^T h_{k,u}^T \sqrt{p_{k,u}^T} s_{k,u}^T}_{\text{Desired Signal}} \\
 & + \underbrace{\sum_{k \in \mathbb{K}} \left(x_{k,u}^T h_{k,u}^T \sum_{j \in \mathbb{U}^T, j \neq u} x_{k,j}^T \sqrt{p_{k,j}^T} s_{k,j}^T \right)}_{\text{Intra-cluster Interference}} + \underbrace{\sum_{j \in \mathbb{U}^S} h_u^S \sqrt{p_j^S} s_j^S}_{\text{Cross-tier Interference}} + n_u^T,
 \end{aligned} \tag{4}$$

where $h_{k,u}^T$ and h_u^S represent the channel coefficient from the BS k to the TU u and the channel coefficient from the LEO to the TU u . $n_{k,u}^T$ is the AWGN of the terrestrial system with $E[n_u^T] = \sigma_u^{T^2}$. For the TUs of \mathbb{M}_k with $\forall k \in \mathbb{K}$, we assume that the channel gains are sort as $|h_{k,1}^T|^2 \leq \dots \leq |h_{k,u}^T|^2 \leq \dots \leq |h_{k,U^T}^T|^2$. Let $\{s_u^k\}$ denote the set from 1st TU to the u th TU served by BS k , and k' denote the SU that use the same SC with BS K . The SINR of the TU u is

$$\gamma_u^T = \frac{\sum_{k \in \mathbb{K}} x_{k,u}^T |h_{k,u}^T|^2 p_{k,u}^T}{\sum_{k \in \mathbb{K}} x_{k,u}^T |h_{k,u}^T|^2 \sum_{j \in \mathbb{M}_k \setminus \{s_u^k\}} p_{k,j}^T + \sum_{k \in \mathbb{K}} x_{k,u}^T |h_u^S|^2 p_{k'}^S + \sigma_u^{T^2}}, \tag{5}$$

where $\sigma_u^{T^2} = B_{sc}N_0^T$. Thus, the transmission rate of the TU u is

$$C_u^T = \log_2(1 + \gamma_u^T). \tag{6}$$

2.2 Stackelberg Game Formulation

In the terrestrial-satellite networks, the resource optimization problem can be formulated as the primary problem for satellite system and the secondary problem for the terrestrial system.

The SUs as the primary user, the QoS should be guaranteed first. On the other hand, the power consumption of the satellite system is considered to reduce waste of limited resources and minimize cross-tier interference. Accordingly, we construct the utility function of each SU as

$$v_u^S = C_u^S - \omega_u^S p_u^S, \tag{7}$$

where ω_u^S is a weighting factor in bits/s/(Hz·W). The sum utility of the satellite system is $V^S = \sum_{u \in \mathbb{U}^S} u_u^S$. Thus, the primary optimization problem for satellite system can be modeled as

$$\begin{aligned}
\mathcal{P}1 : \quad & \max_{\{p_u^S\}} V^S \\
\text{s.t. } C1 : & C_u^S \geq C^{th}, \forall u \in \mathbb{U}^S, \\
C2 : & \sum_{u \in \mathbb{U}^S} p_u^S \leq P_{\max}^S, \\
C3 : & p_u^S \geq 0, \forall u \in \mathbb{U}^S,
\end{aligned} \tag{8}$$

where C1 guarantees the QoS of SUs with the rate threshold. C2 and C3 denote the power constraint of SUs.

According to the power optimization strategy of the satellite-tier, the TU selects the access points and transmission power to maximize its utility function. To achieve the fairness of TUs, we regard the minimum worst-case transmission rate among TUs as the utility function in the terrestrial-tier, i.e., $z = \min_{\forall u \in \mathbb{U}^T} C_u^T$. The secondary problem can be formulated as

$$\begin{aligned}
\mathcal{P}2 : \quad & \max_{\{x_{k,u}^T, p_{k,u}^T\}} z \\
\text{s.t. } C1 : & C_u^T \geq z, \forall u \in \mathbb{U}^T, \\
C2 : & \sum_{i \in \mathbb{M}_k} p_{k,i}^T \leq P_{\max}^T, \forall k \in \mathbb{K}, \\
C3 : & p_{k,u}^T \geq 0, \forall k \in \mathbb{K}, \forall u \in \mathbb{U}^T, \\
C4 : & |g_{k,k'}^T|^2 \sum_{i \in_k} p_{k,i}^T \leq I^{th}, \forall k \in \mathbb{K}, \\
C5 : & x_{k,u}^T \in \{0, 1\}, \forall k \in \mathbb{K}, \forall u \in \mathbb{U}^T, \\
C6 : & \mathbb{M}_k = \{x_{k,i}^T = 1 | i \in \mathbb{U}^T\}, |\mathbb{M}_k| \leq M^{BS}, \forall k \in \mathbb{K}, \\
C7 : & \sum_{k \in \mathbb{K}} x_{k,u}^T \leq M^{UE}, \forall u \in \mathbb{U}^T,
\end{aligned} \tag{9}$$

where C1 is the minimum rate guarantee. C2 and C3 are the power constraint of the TUs. C4 ensures the SU's QoS. C5, C6, and C7 denote that the cooperative factor $x_{k,u}^T$ is a binary variable, and restrict the maximum number for cooperative transmission of BSs and TUs, respectively.

The aforementioned two-tier resource allocation problem can be modeled as a Stackelberg game model, where the competition between leaders and followers is recognized as the utility maximization between SUs and TUs. SUs as the leaders have the strategy space $\mathbf{P}^L = \{\mathbf{P}^S\}$, and TUs as the followers have the strategy space $\mathbf{P}^F = \{\mathbf{X}^T, \mathbf{P}^T\}$, that construct the strategy space $\{\mathbf{P}^L, \mathbf{P}^F\}$ of the game. Leaders choose strategy \mathbf{P}^S firstly in the game. In the follower sub-game, each TU as a player of the non-cooperative game, tries to choose the

strategy \mathbf{P}^F to maximize their transmission rate. Afterwards, the leaders update their strategies based on the strategies of followers and vice versa. According to the Schauder fixed point theory [14] and the theorem that there must be a Nash equilibrium point in finite strategic game. It is easy to prove that both leader game and follower game can reach Nash equilibrium, so the existence of Stackelberg equilibrium point can be achieved.

3 Optimization Solutions of the Stackelberg Game

In this section, we firstly optimize the leader sub-game problem and follower sub-game problem, respectively. Then the joint iterative optimization algorithm is proposed to achieve Stackelberg equilibrium. Furthermore, the corresponding algorithm design is introduced.

3.1 Optimization Strategy of the Leader Sub-Game

The optimization problem $\mathcal{P}1$ is a single-cell power allocation problem. And with the fixed follower strategy of the terrestrial-tier, $\mathcal{P}1$ is strictly convex where the optimal strategy can be obtained by using the KKT optimality conditions. Let $\frac{\partial U_u^S}{\partial p_u^S} \Big|_{\hat{p}_u^S} = 0$, we obtain the extreme point $\hat{p}_u^S = \frac{1}{\ln 2\omega_u^S} - \frac{I_u^S}{|g_u^S|^2}$, where $I_u^S = |g_{k,u}^T|^2 \sum_{j \in \mathbb{M}_k} x_{k,j}^T p_{k,j}^T + \sigma^S$. In the meanwhile, C1 of $\mathcal{P}1$ is equivalently converted into

$$p_u^S \geq \frac{(2^{C^{th}} - 1)I_u^S}{|g_u^S|^2}. \quad (10)$$

Defining $\Phi_u = \frac{(2^{C^{th}} - 1)I_u^S}{|g_u^S|^2}$. Then the optimal strategy of leader sub-game $\mathcal{P}1$ can be expressed as

$$\tilde{p}_u^S = \begin{cases} \Phi_u, & \hat{p}_u^S \leq \Phi_u \\ \frac{\Phi_u}{\sum_{u \in \mathbb{U}^S} \Phi_u} P_{\max}^S, & \hat{p}_u^S \geq \frac{\Phi_u}{\sum_{u \in \mathbb{U}^S} \Phi_u} P_{\max}^S \\ \hat{p}_u^S, & \text{else} \end{cases}. \quad (11)$$

By the optimal strategy of the leader sub-game, the followers in terrestrial system decides and adjusts resource allocation to achieve its maximum utility.

3.2 Optimization Strategy of the Follower Sub-Game

The follower sub-game $\mathcal{P}2$ is a highly non-convex mixed integer programming (MIP) problem, because of the binary factor $x_{k,u}^T$, and the non-convexity constraint of C1, which cannot be solved directly in a polynomial time. In this part, we divide the follower optimization problem into two sub-problems: cooperative scheduling problem and power allocation problem.

The cooperative transmission problem is essentially a two-sided matching problem between TUs and BSs. An evolutionary gale-shapley algorithm is considered in this paper to solve the cooperative transmission problem. Firstly, we introduce a cooperative threshold h^{th} and calculate $w_{k,n}^T = \text{sgn} \left(\left\lfloor \frac{|h_{k,u}^T|^2}{h^{th}} \right\rfloor \right)$, where $w_{k,n}^T = 0$ means that the BS k cannot serve TU u due to the poor channel condition. $\lfloor \cdot \rfloor$ is the rounding down function.

In the evolutionary gale-shapley algorithm, each TU selects the BS in turn according to the preference sequence based on channel coefficient. The BS makes a decision by the value of $|\mathbb{M}_k|$ and the cluster rate. If the BS k rejects the TU u , the TU u will no longer send the request to BS k in the following cycles. The detailed description can be found in **Algorithm 1**.

With the optimization cooperative scheduling strategy \mathbf{X}^T , the max-min problem of user rate optimization is convex with respect to the power allocation $\{p_{k,u}^T\}$. The Lagrangian dual decomposition approach is used to obtain the power solutions. The corresponding Lagrangian function is

$$L(p_{k,u}^T, \lambda_u, \mu_k, \nu_k) = D(p_{k,u}^T, \lambda_u, \mu_u, \nu_k) + z(1 - \sum_{u \in \mathbb{U}^T} \lambda_u) + I^{th} \sum_{k \in \mathbb{K}} \mu_k + P_{\max}^T \sum_{k \in \mathbb{K}} \nu_k, \quad (12)$$

where

$$D(p_{k,u}^T, \lambda_u, \mu_k, \nu_k) = \sum_{u \in \mathbb{U}^T} \lambda_u \log_2 \left(1 + \frac{\sum_{k \in \mathbb{K}} x_{k,u}^T |h_{k,u}^T|^2 p_{k,u}^T}{\sum_{k \in \mathbb{K}} |h_{k,u}^T|^2 \sum_{j \in \mathbb{M}_k \setminus \{s_u^k\}} p_{k,j}^T + I_u^T} \right) - \sum_{k \in \mathbb{K}} \mu_k |g_{k,k'}^T|^2 \sum_{u \in \mathbb{M}_k} p_{k,u}^T - \sum_{k \in \mathbb{K}} \nu_k \sum_{u \in \mathbb{M}_k} p_{k,u}^T, \quad (13)$$

where $I_u^T = \sum_{k \in \mathbb{K}} x_{k,u}^T |h_u^S|^2 p_{k'}^S + \sigma_u^{T^2}$. The optimal power $p_{k,u}^{T*}$ can be derived by

$\frac{\partial L(p_{k,u}^T, \lambda_u, \mu_u, \nu_k)}{\partial p_{k,u}^T} = 0$, which is

$$p_{k,u}^{T*} = \left[\frac{\lambda_u x_{k,u}^T}{(\mu_u |g_{k,k'}^T|^2 + \nu_k) \ln 2} - \frac{\sum_{k \in \mathbb{K}} x_{k,u}^T |h_{k,u}^T|^2 \sum_{j \in \mathbb{M}_k \setminus \{s_u\}} p_{k,j}^T + I_u^T}{|h_{k,u}^T|^2} \right]^+, \quad (14)$$

where $[y]^+ = \max(y, 0)$. Let σ_{λ_u} , σ_{μ_k} , and σ_{ν_k} be the positive step size at the l th iteration, the updated expressions of the dual variables are

$$\lambda_u(l+1) = [\lambda_u(l) - \sigma_{\lambda_u} (C_u^T - z)]^+, \quad (15)$$

$$\mu_k(l+1) = \left[\mu_k(l) - \sigma_{\mu_k} (I^{th} - |g_{k,k'}^T|^2 \sum_{u \in \mathbb{M}_k} p_{k,u}^T) \right]^+, \quad (16)$$

$$\nu_k(l+1) = \left[\nu_k(l) - \sigma_{\nu_k} \left(P_{\max}^T - \sum_{u \in \mathbb{M}_k} p_{k,u}^T \right) \right]^+ . \quad (17)$$

3.3 An Joint Iterative Algorithm to Achieve Stackelberg Equilibrium

According to the aforementioned leader and follower strategies, a joint iterative algorithm is proposed in this section. At each iteration, the SUs choose the optimization strategy by (11), the TUs then find the follower optimization strategy based on the leader strategy. The Stackelberg equilibrium solutions eventually can be obtained when the algorithm reaches convergence. The detailed process is described in **Algorithm 1**.

Algorithm 1. The Stackelberg Equilibrium point for the two-tier TSNs

Input:

Channel coefficient $\{g_u^S\}$, $\{g_{k,u}^T\}$, $\{h_u^S\}$, and $\{h_{k,u}^T\}$;
 Maximum satellite power P_{\max}^S and BS power P_{\max}^T ;
 Rate threshold C^{th} and tolerable interference temperature I^{th} of SUs.

Output:

\mathbf{p}^S , \mathbf{X}^T , \mathbf{p}^T .

- 1: Initialize maximum Stackelberg iterative number T_{\max} and maximum Lagrangian iterative number L_{\max} ; Let $t = 0$;
- 2: Initialize $F_{unmatch} = \mathbb{U}^t$, $w_{k,n}^T$, n_u , and m_k ; Set preference sequence $\{PF_TU_u\}$.
- 3: **repeat**
- 4: **Leader Strategy for Satellite-Tier**
- 5: **for** $u \in \mathbb{U}^S$ **do**
- 6: Calculate and update leader power strategy p_u^S by (11);
- 7: **end for**
- 8: **Follower Strategy for Terrestrial-Tier**
- 9: **while** $F_{unmatch} \neq \emptyset$ **do**
- 10: **for** $u \in F_{unmatch}$ **do**
- 11: TU u sends matching request to BS k with $w_{k,n}^T = 1$ according to $\{PF_TU_u\}$;
- 12: **if** $m_k < M^{BS}$ **then**
- 13: Update \mathbf{X}^T , $n_u = n_u + 1$.
- 14: **else**
- 15: BS k choose the user cluster with the highest cluster rate, update \mathbf{X}^T ;
- 16: $n_u = n_u + 1$ and replaced TU \tilde{u} with $n_{\tilde{u}} = n_{\tilde{u}} - 1$;
- 17: $F_{unmatch} = F_{unmatch} \cup \{\tilde{u}\}$.
- 18: **end if**
- 19: **if** $n_u = M^{UE}$ **then**
- 20: $F_{unmatch} = F_{unmatch} \setminus \{u\}$.
- 21: **end if**
- 22: **end for**
- 23: **end while**
- 24: $l = 0$;
- 25: **while** $l < L_{\max}$ **do**
- 26: **for** $u \in \mathbb{U}^T$, $k \in \mathbb{K}$ **do**

```

27:         Calculate and update follower power strategy  $p_{k,u}^T$  by (14);
28:         Calculate and update Lagrangian dual variables  $\lambda_u$ ,  $\mu_k$ , and  $\nu_k$  by (15), (16), and
           (17), respectively.
29:     end for
30:      $l = l + 1$ ;
31: end while
32:      $t = t + 1$ ;
33: until Coverage or  $t = T_{\max}$ 

```

4 Simulation Results and Discussion

In this section, performance of the proposed algorithm is evaluated. The satellite altitude is 36000 km covering an urban area with 20 terrestrial BSs, and the radius of each cell is 1 km. The carrier frequency of TSN is 3 GHz and the system bandwidth is 10 MHz. The number of SUs and TUs are 20 and 100. The maximum transmission power of satellite $P_{\max}^S = 53$ dBm and the terrestrial BS $P_{\max}^T = 43$ dBm. The terrestrial channel is based on the UMa_A model [15] and Rayleigh fading. And the satellite channel is consider as the shadowed-Rician model [6]. The other transmission parameters are set referring to [16,17].

Figure 1 shows the convergence of several SUs for the proposed algorithm with $C^{th} = 5$ bits/s/Hz and $I^{th} = -60$ dBm. From the Fig. 1, we can see that the curves converge to stable solutions at about 5 Stackelberg iterations. This is because that the solutions of the two sub-games are convergent, thus the joint iteration has a good convergence, which verifies conversely the previous analysis.

Figure 2 shows the convergence of follower sub-game for the proposed algorithm with different C^{th} , where $I^{th} = -60$ dBm. Similar to Fig. 1, the follower strategy converges at about 6 iterations under different C^{th} . With the increase of C^{th} , the followers obtain the lower transmission rate z , because more stringent QoS of SUs shrinks the feasible region of the follower strategy \mathbf{P}^t to reduce cross-tier interference, resulting in the low transmission rate of TUs.

Figure 3(a) and Fig. 3(b) show the minimum transmission rate of TUs and system sum rate when the number of TUs increases from 10 to 60, respectively, where $C^{th} = 5$ bits/s/Hz and $I^{th} = -40$ dBm. It can be seen that the system sum rate of TSN increases along with an increase in the number of TUs, while the minimum transmission rate of terrestrial-tier decreases accordingly. The reason for this trend is that the lower power ratio is achieved for each TUs with the limited P_{\max}^T as the number of TUs increases, that results to a decreased transmission rate.

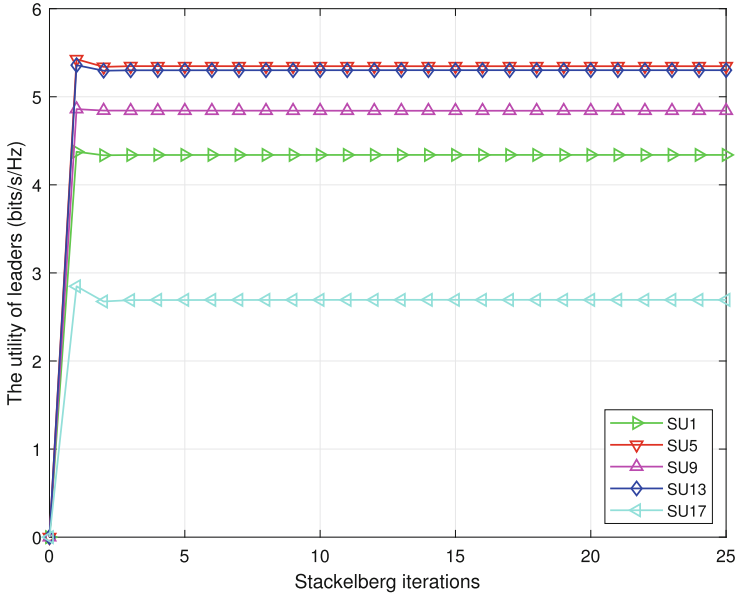


Fig. 1. Utility of leaders versus the number of Stackelberg iteration.

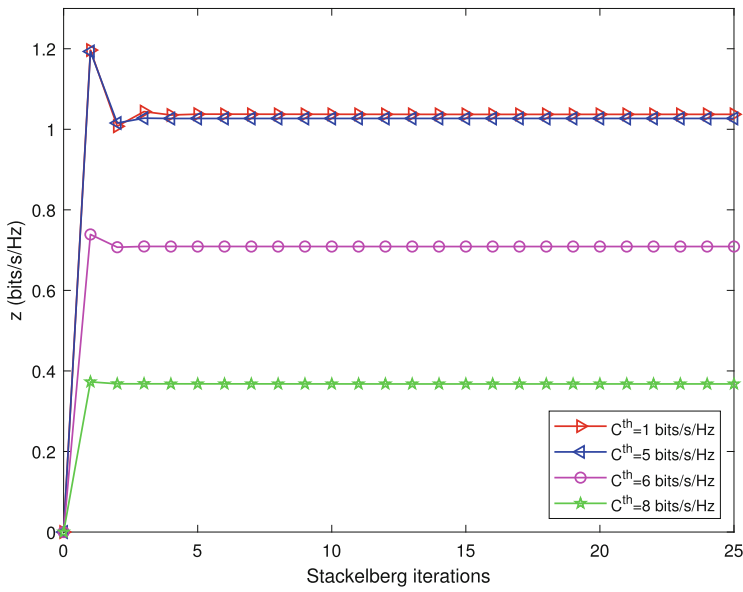


Fig. 2. Minimum rate z versus the number of Stackelberg iterations with different C^th .

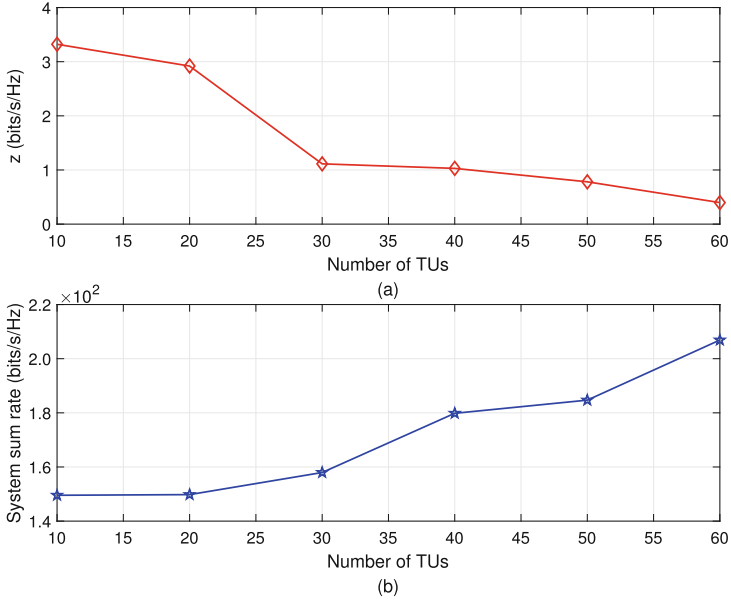


Fig. 3. (a) Minimum rate z versus the number of TUs. (b) System sum rate versus the number of TUs.

5 Conclusion

In this paper, we discussed the interference management problem of the two-tier TSNs by considering QoS guarantee, user fairness, cooperative scheduling, and power allocation. The two-tier optimization problem can be formulated as a Stackelberg game model where the SUs were the leaders and the TUs were the followers. The leader strategy was obtained by maximizing the SU utility. The followers aimed to the minimum transmission rate maximization problem to achieve the fairness among the TUs. To solve this highly non-convex MIP problem, we divided it into the cooperative scheduling problem and power allocation problem. An evolutionary gale-shapley algorithm was proposed to solve the cooperative scheduling problem. Accordingly, the power allocation problem can be transformed as a convex one, and the closed-form expression was derived. Finally, a joint iterative algorithm was proposed to achieve Stackelberg Equilibrium between the leader strategy and the follower strategy. The simulation results showed that the proposed algorithm has a good convergence and performance improvement.

References

1. Takahashi, M., Kawamoto, Y., Kato, N., Miura, A., Toyoshima, M.: DBF-based fusion control of transmit power and beam directivity for flexible resource alloca-

- tion in HTS communication system toward B5G. *IEEE Trans. Wireless Commun.* **21**(1), 95–105 (2022)
2. Fu, S., Gao, J., Zhao, L.: Collaborative multi-resource allocation in terrestrial-satellite network towards 6G. *IEEE Trans. Wirel. Commun.* **20**(11), 7057–7071 (2021)
 3. Wu, B., Fang, F., Fu, S.: Improving the system performance in terrestrial-satellite relay networks by configuring aerial relay. *IEEE Trans. Veh. Technol.* **70**(12), 13139–13148 (2021)
 4. Shi, Y., Liu, J., Fadlullah, Z.M., Kato, N.: Cross-layer data delivery in satellite-aerial-terrestrial communication. *IEEE Wirel. Commun.* **25**(3), 138–143 (2018)
 5. Zhang, Y., Zhang, H., Zhou, H., Long, K., Karagiannidis, G.K.: Resource allocation in terrestrial-satellite based next generation multiple access networks with interference cooperation. *IEEE J. Sel. Areas Commun.* **40**(4), 1210–1221 (2022)
 6. Ruan, Y., Li, Y., Wang, C.-X., Zhang, R., Zhang, H.: Energy efficient power allocation for delay constrained cognitive satellite terrestrial networks under interference constraints. *IEEE Trans. Wireless Commun.* **18**(10), 4957–4969 (2019)
 7. Liu, X., Zhai, X.B., Lu, W., Wu, C.: QoS-guarantee resource allocation for multi-beam satellite industrial internet of things with NOMA. *IEEE Trans. Industr. Inf.* **17**(3), 2052–2061 (2021)
 8. Peng, D., Bandi, A., Li, Y., Chatzinotas, S., Ottersten, B.: Hybrid beamforming, user scheduling, and resource allocation for integrated terrestrial-satellite communication. *IEEE Trans. Veh. Technol.* **70**(9), 8868–8882 (2021)
 9. Fang, F., Xu, Y., Ding, Z., Shen, C., Peng, M., Karagiannidis, G.K.: Optimal resource allocation for delay minimization in NOMA-MEC networks. *IEEE Trans. Commun.* **68**(12), 7867–7881 (2020)
 10. Xu, Y., et al.: Robust resource allocation for two-tier HetNets: an interference-efficiency perspective. *IEEE Trans. Green Commun. Netw.* **5**(3), 1514–1528 (2021)
 11. Zhu, X., Jiang, C., Kuang, L., Zhao, Z., Guo, S.: Two-layer game based resource allocation in cloud based integrated terrestrial-satellite networks. *IEEE Trans. Cogn. Commun. Netw.* **6**(2), 509–522 (2020)
 12. Abdu, T.S., Kisseleff, S., Lagunas, E., Chatzinotas, S.: Flexible resource optimization for GEO multibeam satellite communication system. *IEEE Trans. Wireless Commun.* **20**(12), 7888–7902 (2021)
 13. Ali, M.S., Hossain, E., Al-Dweik, A., Kim, D.I.: Downlink power allocation for CoMP-NOMA in multi-cell networks. *IEEE Trans. Commun.* **66**(9), 3982–3998 (2018)
 14. Talman, A.J.J.: Variable dimension fixed point algorithms and triangulations. *Stat. Neerl.* **35**(1), 1 (1981)
 15. Guidelines for evaluation of radio interface technologies for IMT-2020, Report ITU-R M.2412-0, ITU (2017)
 16. Li, J., Xue, K., Wei, D.S.L., Liu, J., Zhang, Y.: Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks. *IEEE Trans. Wireless Commun.* **19**(4), 2367–2381 (2020)
 17. Miller, M.J., Vucetic, B., Berry, L.: *Satellite Communications: Mobile and Fixed Services*. Springer, Berlin (1993)



Optimal Resource Allocation for Computation Offloading in Maritime Communication Networks: An Energy-Efficient Design via Matching Game

Minghui Dai¹, Zhishen Luo¹, Tianshun Wang¹, Yuan Wu^{1,2(✉)}, Liping Qian³, and Bin Lin⁴

¹ State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau, China

{minghuidai, mc14891, YC07478, yuanwu}@um.edu.mo

² Zhuhai UM Science and Technology Research Institute, Zhuhai, China

³ College of Information Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang, China
lpqian@zjut.edu.cn

⁴ Department of Communication Engineering, Dalian Maritime University, Dalian, China
binlin@dlmu.edu.cn

Abstract. The increasing growth of maritime activities leads to the challenges for processing the maritime data. However, the resources-limited maritime devices cannot meet the requirements of transmission delay and energy consumption. In this paper, we investigate the resource allocation for computation offloading in maritime communication networks via game theory to improve the offloading efficiency and reduce the energy consumption of maritime devices. Specifically, we propose an optimization problem that jointly optimizes the offloading data, the computation resource allocation of unmanned surface vehicle (USV) and the allocation of acoustic channels, with the objective of minimizing the total energy consumption of underwater wireless sensor (UWS). Despite the non-convexity of the joint optimization problem, we propose a layered structure and decompose it into a top-problem for optimizing the data offloading, a middle-problem for optimizing the computation resource allocation of USV, a bottom problem for optimizing the channel allocation. We conduct simulations to validate the effectiveness and efficiency of the proposed algorithms.

This work was supported in part by National Natural Science Foundation of China under Grants 62122069, 62072490, and 62071431, in part by FDCT-MOST Joint Project under Grant 0066/2019/AMJ, in part by Science and Technology Development Fund of Macau SAR under Grants 0060/2019/A1 and 0162/2019/A3, in part by the Intergovernmental International Cooperation in Science and Technology Innovation Program under Grant 2019YFE0111600, in part by FDCT SKL-IOTSC(UM)-2021-2023, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515011287 and GDST 2020B1212030003, and in part by Research Grant of University of Macau under Grants MYRG2020-00107-IOTSC and MYRG2018-00237-FST.

Keywords: Maritime communication networks · Computation offloading · Energy efficiency

1 Introduction

Maritime communication networks, which are developed to facilitate ocean data gathering, resource exploration, and environment monitoring, have attracted a growing attention in recent years [1–3]. However, the harsh maritime environment and the high cost of underwater equipment pose critical constraints to construct the development of maritime applications. The typical characteristic of maritime communication is the underwater transmission, which is commonly associated with low data rate and high data packet loss due to the complicated maritime environment [4,5]. Therefore, it is important to investigate the underwater channel allocation problem for underwater acoustic communication with the objective of improving the communication efficiency [6,7].

To collect the ocean information, a group of underwater wireless sensors (UWSs) are usually deployed on the seabed to perceive ocean data. However, these UWSs are generally battery-powered [8,9], which are hard to recharge due to the harsh ocean environment. Therefore, the data collected by UWSs needs to be transmitted to the sea surface nodes for further processing. Maritime edge computing is a promising paradigm to improve the computation efficiency and save the energy consumption of UWSs [10–12]. In particular, unmanned surface vehicles (USVs) sailing on the sea surface can provide edge computation and communication capacities to offer services for UWSs. Besides, game theory is widely used for improving the network resource utilization through designing efficient scheme [13], which is a feasible approach in resource-limited maritime networks. However, due to the complicated underwater communications, it is still a critical issue to design an energy-efficient computation offloading scheme in maritime communication networks for reducing the energy consumption of UWSs.

Considering the limited communication resources in maritime networks, the underwater acoustic communication should be well designed to improve the network throughput. In [14], Yang *et al.* proposed a computation task offloading to minimize the energy consumption of vessel terminals and the execution delay of computation task. In [15], Ma *et al.* presented an unmanned aerial vehicle (UAV) aided ocean monitoring network for remote oceanic data collection to maximize the lifetime of UWSs. In [16], Yang *et al.* investigated an offloading optimization problem and proposed a two-stage joint optimal offloading algorithm for optimizing communication resource allocation under limited energy and sensitive latency. Different from the above works, we investigate an optimal resource allocation for computation offloading in maritime communication networks in which a group of UWSs are deployed at the seabed to monitor the ocean environment. A USV sails on the sea to collect data and provide computation offloading service for UWSs. We aim to minimize the energy consumption of UWSs by jointly optimizing the acoustic channel allocation, data offloading, and computation resource allocation of USV. The main contributions of this paper can be summarized as follows:

- Considering the resources-limited maritime devices, we investigate the energy efficient computation offloading scheme for UWS in maritime communication networks to minimize the energy consumption of UWSs.

- We propose a joint optimization problem for optimizing the allocation of underwater acoustic channels, the offloading data, and the computation resource allocation of USV. Despite the non-convexity of the joint optimization problem, we propose a layered structure to decompose it into a top-problem, a middle problem, and a bottom problem to solve it. Simulation results have demonstrated the effectiveness of the proposed algorithms.

The remainder of this paper is organized as follows. Section 2 presents the system model and problem formulation. Section 3 introduces the proposed algorithms. Simulation performance is provided in Sect. 4, and Sect. 5 concludes this work and discusses the future direction.

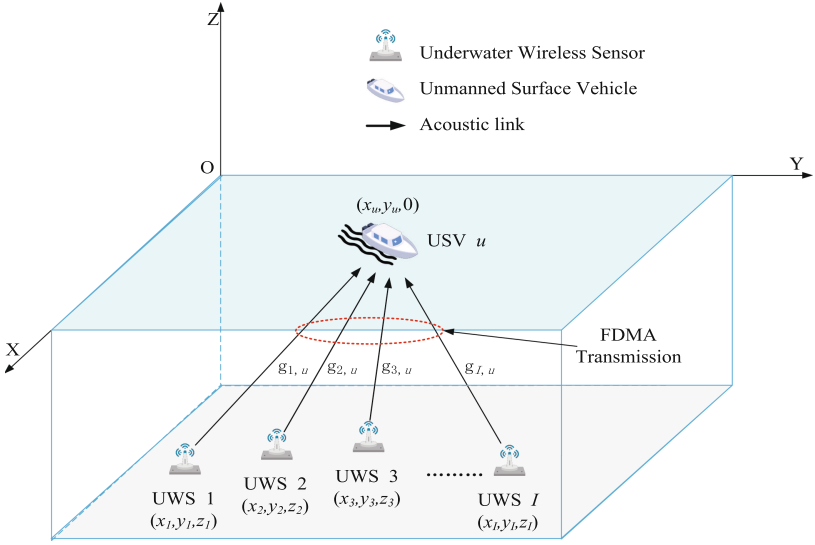


Fig. 1. System model of computation offloading in maritime communication networks.

2 System Model and Problem Formulation

Figure 1 shows the scenario of computation offloading in maritime communication networks. UWSs are deployed at the seabed to monitor the ocean environment, and the set of UWSs is denoted as $\mathcal{I} = \{1, \dots, i, \dots, I\}$. The total workloads of UWS i is denoted as S_i^{tot} . One USV u is deployed on the sea to collect and process the data from UWSs. The underwater acoustic channels are expressed as $\mathcal{K} = \{1, \dots, k, \dots, K\}$. The data transmission in underwater environment is carried over the acoustic signal with a very fast decaying rate, which thus is different from radio transmission over the air. Based on the Urlick's model [17], we can derive the attenuation of underwater acoustic signals, which depends on the transmission distance and the carrier frequency, as follows

$$A(d_{i,u}, k) = (d_{i,u})^\lambda \alpha(k)^{\frac{d_{i,u}}{1000}}, \forall i \in \mathcal{I}, \forall k \in \mathcal{K}, \quad (1)$$

where $d_{i,u}$ denotes the distance between USV u and UWS i . In Eq. (1), k is the central frequency of the acoustic signal. λ is constant which is usually set as 1.5. $\alpha(k)$ is the absorption coefficient. For the convenience of expression, we consider a three-dimensional Cartesian coordinate system for the positions of USV and UWSs. The position of UWS i is denoted as (x_i, y_i, z_i) . The position of USV is denoted as $(x_u, y_u, 0)$. Therefore, we can express the distance between UWS i and USV u as

$$d_{i,u} = \sqrt{(x_u - x_i)^2 + (y_u - y_i)^2 + z_i^2}, \forall i \in \mathcal{I}. \quad (2)$$

The absorption coefficient $\alpha(k)$ can be calculated according to Thorp's empirical formula [17] as

$$10\log_{10}\alpha(k) = \frac{0.11k^2}{1+k^2} + \frac{44k^2}{4000+k^2} + 2.75 \times 10^{-4}k^2 + 0.003, \forall k \in \mathcal{K}. \quad (3)$$

In addition to the path loss, the underwater acoustic signals are also affected by the oceanic noise N_0 and the channel bandwidth W . Therefore, the underwater acoustic channel gain between USV u and UWS i can be expressed as

$$g_i = \frac{1}{A(d_{i,u}, k)N_0W}, \forall i \in \mathcal{I}, \forall k \in \mathcal{K}. \quad (4)$$

Let $a_{i,k} \in \{0, 1\}, \forall i \in \mathcal{I}, k \in \mathcal{K}$ denote the association between underwater acoustic channel k and UWS i . $a_{i,k} = 1$ indicates that channel k is allocated to UWS i . Otherwise, $a_{i,k} = 0$. Then, we can express the received signal to interference plus noise ratio (SINR) of UWS i as

$$\phi_i = \frac{a_{i,k}p_i g_i}{N_0W}, \forall i \in \mathcal{I}, \forall k \in \mathcal{K}, \quad (5)$$

where p_i is the transmission power of UWS i . The transmission rate between UWS i and USV u can be calculated by

$$r_i = W\log_2(1 + \phi_i), \forall i \in \mathcal{I}, \forall k \in \mathcal{K}. \quad (6)$$

Let s_i denote the offloading data of UWS i , we can obtain

$$\frac{s_i}{t_i} = W\log_2(1 + \phi_i), \forall i \in \mathcal{I}, \forall k \in \mathcal{K}, \quad (7)$$

where t_i is the transmission time of UWS i . The required transmission power for UWS i can be expressed as [18]

$$p_i = \frac{N_0W}{a_{i,k}g_{i,k}} \left(2^{\frac{s_i}{Wt_i}} - 1 \right), \forall i \in \mathcal{I}, \forall k \in \mathcal{K}. \quad (8)$$

The energy consumption of UWS i for offloading data can be expressed by

$$E_i = p_i t_i, \forall i \in \mathcal{I}, \forall k \in \mathcal{K}. \quad (9)$$

The local computing time of UWS i can be expressed as

$$t_i^{\text{loc}} = c_i \frac{S_i^{\text{tot}} - s_i}{\mu_i^{\text{loc}}}, \forall i \in \mathcal{I}, \quad (10)$$

where c_i is the number of CPU cycles for processing one bit of data. μ_i^{loc} denotes the processing capability of UWS i which is measured by the CPU cycles per second. Therefore, the energy consumption of local computation is expressed as [19]

$$E_i^{\text{loc}} = \rho_i t_i^{\text{loc}} (\mu_i^{\text{loc}})^3 = c_i (S_i^{\text{tot}} - s_i) \rho_i (\mu_i^{\text{loc}})^2, \forall i \in \mathcal{I}, \quad (11)$$

where ρ_i denotes the power coefficient of UWS i .

Let μ_i^{all} denote the allocated processing capability of USV for UWS i in CPU cycles per second. c_u expresses the number of CPU cycles for USV to process one bit of data. The overall delay for UWS i to complete its total computation workload S_i^{tot} can be expressed as

$$t_i^{\text{ove}} = \max \left\{ c_i \frac{S_i^{\text{tot}} - s_i}{\mu_i^{\text{loc}}}, t_i + c_u \frac{s_i}{\mu_i^{\text{all}}} \right\}, \forall i \in \mathcal{I}. \quad (12)$$

The total energy consumption of UWS i can be calculated by

$$\begin{aligned} E_i^{\text{tot}} &= E_i + E_i^{\text{loc}} \\ &= c_i (S_i^{\text{tot}} - s_i) \rho_i (\mu_i^{\text{loc}})^2 + \frac{N_0 W}{a_{i,k} g_{i,k}} \left(2^{\frac{s_i}{W t_i}} - 1 \right) t_i, \forall i \in \mathcal{I}. \end{aligned} \quad (13)$$

2.1 Problem Formulation

Based on the above modeling, we formulate the following optimization problem that aims to minimize the total energy consumption (TEC) of UWS, i.e.,

$$\text{(TEC)} : \min \left\{ \max_{i \in \mathcal{I}} \{ E_i^{\text{tot}} \} \right\}$$

$$\text{subject to : } \sum_{k=1}^N a_{i,k} = 1, \forall i \in \mathcal{I}, \quad (14)$$

$$\sum_{i=1}^I a_{i,k} \leq 1, \forall k \in \mathcal{K}, \quad (15)$$

$$\sum_{i=1}^I \mu_i^{\text{all}} \leq \mu^{\text{max}}, \quad (16)$$

$$\max \left\{ c_i \frac{S_i^{\text{tot}} - s_i}{\mu_i^{\text{loc}}}, t_i + c_u \frac{s_i}{\mu_i^{\text{all}}} \right\} \leq T_i^{\text{max}}, \forall i \in \mathcal{I}, \quad (17)$$

$$0 \leq s_i \leq S_i^{\text{tot}}, \forall i \in \mathcal{I}, \quad (18)$$

$$0 \leq \frac{N_0 W}{a_{i,k} g_{i,k}} \left(2^{\frac{s_i}{W t_i}} - 1 \right) \leq P_i^{\text{max}}, \forall i \in \mathcal{I}, \forall k \in \mathcal{K}, \quad (19)$$

$$\text{variables : } \{s_i\}_{i \in \mathcal{I}}, \{a_{i,k}\}_{i \in \mathcal{I}, k \in \mathcal{K}}, \{\mu_i^{\text{all}}\}_{i \in \mathcal{I}}.$$

In Problem (TEC), constraints (14) and (15) ensure that each UWS i can only occupy one acoustic channel, and each acoustic channel can only be assigned to one

UWS. Constraint (16) guarantees that the allocated computing-rate of USV cannot exceed the maximum computing-rate μ^{\max} . Constraint (17) ensures that the transmissions delay of each UWS cannot beyond its allowed maximum transmission delay T_i^{\max} . Constraint (18) guarantees that the offloading computation workload of each UWS should below its total computation workload. Constraint (19) indicates that the transmission power of UWS cannot exceed the maximum power P_i^{\max} .

3 Proposed Algorithm for Solving Problem (TEC)

3.1 Layer Structure of Problem (TEC)

Problem (TEC) is a strictly nonconvex optimization problem. The total energy consumption of UWS i is decreasing with the increase of t_i . Based on constraint (17), we consider that the constraint is strictly binding, and we can obtain $t_i = T_i^{\max} - c_u \frac{s_i}{\mu_i^{\text{all}}}$. The total energy consumption of UWS i can be rewritten as

$$E_i^{\text{tot}} = c_i (S_i^{\text{tot}} - s_i) \rho_i (\mu_i^{\text{loc}})^2 + \frac{N_0 W}{a_{i,k} g_{i,k}} \left(2^{\frac{s_i}{W(T_i^{\max} - c_u \frac{s_i}{\mu_i^{\text{all}}})}} - 1 \right) \left(T_i^{\max} - c_u \frac{s_i}{\mu_i^{\text{all}}} \right), \forall i \in \mathcal{I}. \quad (20)$$

Let $\theta = \max_{i \in \mathcal{I}} \{E_i^{\text{tot}}\}$, we obtain $E_i^{\text{tot}} \leq \theta, \forall i \in \mathcal{I}$. Problem (TEC) can be decomposed as the following layer structure

1) Top-problem to find s_i :

$$\begin{aligned} \text{(TEC-Top)} : & \min \theta \\ \text{subject to : } & E_i^{\text{tot}} \leq \theta, \forall i \in \mathcal{I}, \\ & \text{constraints (17), (18), and (19),} \\ \text{variables : } & \{s_i\}_{i \in \mathcal{I}}. \end{aligned} \quad (21)$$

2) Middle-problem to find μ_i^{all} under given $\{s_i\}_{i \in \mathcal{I}}$:

$$\begin{aligned} \text{(TEC-Mid)} : & \min \theta \\ \text{subject to : } & \text{constraint (16) and (21),} \\ \text{variables : } & \{\mu_i^{\text{all}}\}_{i \in \mathcal{I}}. \end{aligned}$$

3) Bottom-problem to find $a_{i,k}$ under given $\{s_i\}_{i \in \mathcal{I}}$ and $\{\mu_i^{\text{all}}\}_{i \in \mathcal{I}}$:

$$\begin{aligned} \text{(TEC-Bot)} : & \min \theta \\ \text{subject to : } & \text{constraints (14), (15), and (21),} \\ \text{variables : } & \{a_{i,k}\}_{i \in \mathcal{I}, k \in \mathcal{K}}. \end{aligned}$$

3.2 Matching Game for Solving Channel Allocation in Problem (TEC-Bot)

In Problem (TEC-Bot), we propose the matching game to solve the channel allocation problem. Specifically, in the channel allocation process, UWSs have different preferences (e.g., channel quality) for acoustic channels. The acoustic channels can rank the UWSs based on the energy efficiency of UWS. Therefore, the matching game can be adopted to model the two-side allocation between UWSs and acoustic channels. Each UWS i is matched to at most one acoustic channel, and each acoustic channel k is allocated to at most one UWS.

The matching game between UWSs and acoustic channels can be described as a tuple $\mathcal{L} = \{\mathcal{I}, \mathcal{K}, \mathcal{P}\}$, where \mathcal{I} and \mathcal{K} are the set of players (i.e., UWSs and acoustic channels), and $\mathcal{I} \cap \mathcal{K} = \emptyset$. \mathcal{P} indicates the preference lists of UWSs and acoustic channels. Let \succ denote the preference relations of UWSs and acoustic channels. $\mathcal{P}(i) |_{\mathcal{L}(i)=k} > \mathcal{P}(i) |_{\mathcal{L}(i)=k'}$, $k \succ k'$ means that UWS i prefers channel k to channel k' . $\mathcal{P}(k) |_{\mathcal{L}(k)=i} > \mathcal{P}(k) |_{\mathcal{L}(k)=i'}$, $i \succ i'$ indicates that channel k prefers UWS i to UWS i' .

Definition 1. Given two disjoint sets of players \mathcal{I} and \mathcal{K} , the matching game for channel allocation is defined as $\mathcal{L}(i) = \{k \in \mathcal{K} | (i, k) \in \mathcal{L}\}$ and $\mathcal{L}(k) = \{i \in \mathcal{I} | (i, k) \in \mathcal{L}\}$, such that (i) $\sum_{k=1}^N a_{i,k} = 1$ and (ii) $\sum_{i=1}^I a_{i,k} \leq 1$.

Note that Definition 1 guarantees that each player on one side matches to a set of players. Constraints (i) and (ii) make sure that the conditions (14) and (15) in Problem (TEC) are satisfied.

The preference relation is utilized to measure the preference levels of each player. Assume that the current matching result is \mathcal{L} and the associated channel allocation vector is represented as $\mathbf{K}^{\mathcal{L}}$. The preference relation of UWS i is defined as the achievable rate when matching with a channel, which can be expressed as

$$\mathcal{U}_i(k) = W \log_2 \left(1 + \frac{P_i g_{i,k}}{N_0 W} \right), \forall i \in \mathcal{I}, \forall k \in \mathcal{K}. \quad (22)$$

For each acoustic channel, let E_i^{\max} denote the maximum energy of UWS i . The preference relation of channel k is defined as the energy efficiency when matching with a UWS, i.e.,

$$\mathcal{U}_k(i) = \log_2 \left(1 + \frac{E_i^{\max}}{E_i^{\text{tot}}} \right), \forall i \in \mathcal{I}. \quad (23)$$

Therefore, for UWS i , the preference relation \succ_i for two channels $k, k' \in \mathcal{K}$, $k \neq k'$ is defined as $k \succ_i k'$ if and only if $\mathcal{U}_i(k) > \mathcal{U}_i(k')$. For channel k , the preference relation \succ_k for two UWSs $i, i' \in \mathcal{I}$, $i \neq i'$ is defined as $i \succ_k i'$ if and only if $\mathcal{U}_k(i) > \mathcal{U}_k(i')$.

Definition 2. In the matching \mathcal{L} , UWS i proposes to match its most preferred channel based on the preference list (i.e., $\mathcal{U}_i(k)$). Channel k proposes to match its most preferred channel based on the preference list (i.e., $\mathcal{U}_k(i)$).

Algorithm 1 presents the proposed channel allocation scheme based on matching game. The key steps are described as follows:

- *Step 4 to Step 9*: UWS $i \in \mathcal{I}$ first searches for the most preferred channel. If a channel k receives more than one proposal from UWSs, denoted by \mathcal{D} , it matches with the most preferred UWS $i \in \mathcal{D}$. Otherwise, it matches with the current requested UWS.
- *Step 10 to Step 14*: The current matching players, i.e., i and k , are removed from the sets \mathcal{I} and \mathcal{K} . The matching process is conducted until there exist no UWSs blocking the current matching.

Algorithm 1. Channel Allocation Scheme based on Matching Game

- 1: Input: the set of players \mathcal{I} and \mathcal{K} .
 - 2: Initialize: compute the preference relations of UWS i and channel k using Eqs. (22) and (23).
 - 3: **while** there exist UWSs blocking the current matching **do**
 - 4: **for** any UWS $i \in \mathcal{I}$ **do**
 - 5: It searches for the most preferred channel k .
 - 6: **if** channel k receives more than one proposal from UWSs in \mathcal{D} **then**
 - 7: It matches with the most preferred UWS $i \in \mathcal{D}$.
 - 8: **else**
 - 9: It matches with the current requested UWS.
 - 10: **end if**
 - 11: Remove the current matching players.
 - 12: **end for**
 - 13: **end while**
 - 14: Output: the matching result $\mathcal{L}(i) = \{k \in \mathcal{K} \mid (i, k \in \mathcal{L})\}$.
-

Proposition 1. *The proposed channel allocation scheme based on matching game is stable. For any UWS $i \in \mathcal{I}$, there exist no UWSs blocking the current matching.*

Proof. The contradiction method is used to prove the stability of the proposed matching algorithm. Specifically, assume that UWS $i \in \mathcal{I}$ and channel $k \in \mathcal{K}$ have no matching after the final matching process. However, both UWS $i \in \mathcal{I}$ and channel $k \in \mathcal{K}$ prefer to match with each other, which means that the matching result $\mathcal{L}(i) = \{k \in \mathcal{K} \mid (i, k \in \mathcal{L})\}$ is restricted. Considering that both UWS $i \in \mathcal{I}$ and channel $k \in \mathcal{K}$ prefer to match with each other, the UWS i must have proposed to match with channel k in the matching process, and the channel $k \in \mathcal{K}$ has rejected the proposal of UWS i , which contradicts the matching $\mathcal{L}(k) = \{i \in \mathcal{I} \mid (i, k \in \mathcal{L})\}$. Therefore, the proposed channel allocation scheme based on matching game can achieve a stable matching.

3.3 Analysis of Problem (TEC-Mid) for Solving μ_i^{all}

Proposition 2. *Given $\{s_i\}_{i \in \mathcal{I}}$ and $\{a_{i,k}\}_{i \in \mathcal{I}, k \in \mathcal{K}}$, Problem (TEC-Mid) is a strictly convex problem.*

Proof. In Problem (TEC-Mid), the first derivative of E_i^{tot} with respect to t_i can be expressed as

$$\frac{\partial E_i^{tot}}{\partial t_i} = \frac{N_0 \left(W t_i \left(-1 + 2^{\frac{s_i}{W t_i}} \right) - 2^{\frac{s_i}{W t_i}} s_i \ln 2 \right)}{t_i a_{i,k} g_{i,k}}. \quad (24)$$

The second derivative of E_i^{tot} with respect to t_i can be expressed as

$$\frac{\partial^2 E_i^{tot}}{\partial t_i^2} = \frac{(\ln 2)^2 2^{\frac{s_i}{W t_i}} N_0 (s_i)^2}{W t_i^3 a_{i,k} g_{i,k}} > 0. \quad (25)$$

With (25), we find that E_i^{tot} is strictly convex with respect to t_i , and the first derivative $\frac{\partial E_i^{tot}}{\partial t_i}$ increases with t_i . Due to $\lim_{t_i \rightarrow \infty} \frac{\partial E_i^{tot}}{\partial t_i} < 0$, E_i^{tot} is convex and non-increasing with respect to t_i . Furthermore, t_i is strictly concave with respect to μ_i^{all} . Therefore, E_i^{tot} is strictly convex with respect to μ_i^{all} .

Proposition 2 indicates that E_i^{tot} is decreasing with μ_i^{all} . With constraint (21), we can derive $\mu_i^{all} \geq \ell_i(\theta)$. With constraint (16), we can obtain $\sum_{i=1}^I \ell_i(\theta) \leq \mu^{\max}$. Therefore, we can find the optimal value of μ_i^{all} by searching the value of θ . Algorithm 2 presents the iterative searching algorithm for solving Problem (TEC-Mid). The key steps are as follows: we first set the upper bound value as $\theta = \theta^{\max}$. For any UWS $i \in \mathcal{I}$, we need to evaluate the conditions $\mu_i^{all} \geq \ell_i(\theta)$ and $\sum_{i=1}^I \ell_i(\theta) \leq \mu^{\max}$. If it is satisfied, the current θ is feasible, and we update the value of θ with a decreasing way. Otherwise, Problem (TEC-Mid) is infeasible.

Algorithm 2. Iterative Searching Algorithm for Solving Problem (TEC-Mid)

- 1: Input: the iterative step size Δ .
 - 2: Initialize: set the upper bound value as $\theta = \theta^{\max}$.
 - 3: **while** $\sum_{i=1}^I \ell_i(\theta) \leq \mu^{\max}$ **do**
 - 4: **for** any UWS $i \in \mathcal{I}$ **do**
 - 5: **if** $\mu_i^{all} \geq \ell_i(\theta)$ **then**
 - 6: Current θ is feasible for Problem (TEC-Mid).
 - 7: Update $\theta \leftarrow \theta - \Delta$.
 - 8: **else**
 - 9: Problem (TEC-Mid) is infeasible.
 - 10: **end if**
 - 11: **end for**
 - 12: **end while**
 - 13: Output: the minimum θ and the corresponding E_i^{tot} .
-

3.4 Analysis of Problem (TEC-Top) for Solving s_i

Proposition 3. Given $\{a_{i,k}\}_{i \in \mathcal{I}, k \in \mathcal{K}}$ and $\{\mu_i^{all}\}_{i \in \mathcal{I}}$, Problem (TEC-Top) is a strictly convex problem.

Proof. In Problem (TEC-Top), the first derivative of E_i^{tot} with respect to s_i can be expressed as

$$\frac{\partial E_i^{\text{tot}}}{\partial s_i} = \frac{N_0 W \left(c_u \left(1 - 2^{\frac{s_i}{T_i^{\text{max}} - c_u \frac{s_i}{\mu_i^{\text{all}} \Pi}} \right) + \frac{2^{\frac{s_i}{T_i^{\text{max}} - c_u \frac{s_i}{\mu_i^{\text{all}} \Pi}} T_i^{\text{max}} \ln 2 (\mu_i^{\text{all}})^2}{W(-c_u s_i + T_i^{\text{max}} c)} \right)}{\mu_i^{\text{all}} a_{i,k} g_{i,k}} - c_i \rho_i (\mu_i^{\text{loc}})^2. \quad (26)$$

The second derivative of E_i^{tot} with respect to s_i can be expressed as

$$\frac{\partial^2 E_i^{\text{tot}}}{\partial s_i^2} = \frac{(\ln 2)^2 2^{\frac{s_i}{T_i^{\text{max}} - c_u \frac{s_i}{\mu_i^{\text{all}} \Pi}} (T_i^{\text{max}})^2 N_0 W (\mu_i^{\text{all}})^3}{W(-c_u s_i + T_i^{\text{max}} \mu_i^{\text{all}})^3 a_{i,k} g_{i,k}}. \quad (27)$$

With constraint (17), we can obtain $T_i^{\text{max}} \mu_i^{\text{all}} - c_u s_i > 0$. Therefore, the value of $\frac{\partial^2 E_i^{\text{tot}}}{\partial s_i^2}$ is always higher than zero. The objective function of E_i^{tot} is strictly convex with respect to s_i .

The stagnation point can be obtained by $\frac{\partial E_i^{\text{tot}}}{\partial s_i} = 0$, which is denoted by s_i^* . To obtain the minimum value of E_i^{tot} , we need to analyze the feasible region of s_i . Specifically, with (17), we can equivalently obtain $S_i^{\text{tot}} - \frac{\mu_i^{\text{loc}} T_i^{\text{max}}}{c_i} \leq s_i$. The lower bound is $s_i^{\text{low}} = S_i^{\text{tot}} - \frac{\mu_i^{\text{loc}} T_i^{\text{max}}}{c_i}$. With (19), we can obtain the constraint $s_i \leq \frac{r_{i,k}^{\text{max}} \mu_i^{\text{all}} T_i^{\text{max}}}{\mu_i^{\text{all}} + r_{i,k}^{\text{max}}}$. Further combining with (18), we can obtain the upper bound of s_i as $s_i^{\text{upp}} = \min \left\{ S_i^{\text{tot}}, \frac{r_{i,k}^{\text{max}} \mu_i^{\text{all}} T_i^{\text{max}}}{\mu_i^{\text{all}} + r_{i,k}^{\text{max}}} \right\}$. Therefore, the feasible region of s_i is $[s_i^{\text{low}}, s_i^{\text{upp}}]$. There are three following cases to be discussed:

- *Case 1:* The stagnation point s_i^* satisfies $s_i^{\text{low}} < s_i^* < s_i^{\text{upp}}$. The optimal value for minimizing E_i^{tot} is obtained at s_i^* .
- *Case 2:* The stagnation point s_i^* satisfies $s_i \leq s_i^{\text{low}}$. The optimal value for minimizing E_i^{tot} is obtained at s_i^{low} .
- *Case 3:* The stagnation point s_i^* satisfies $s_i^{\text{upp}} \leq s_i$. The optimal value for minimizing E_i^{tot} is obtained at s_i^{upp} .

4 Simulation Performance

This section provides the simulation results to validate the effectiveness of the proposed algorithms. We consider a simulation scenario with a size of 1 km \times 1 km \times 0.5 km to model the underwater environment. One USV is deployed on the upper surface and four UWSs are deployed on the seabed to collect ocean data. There are four acoustic channels, and the bandwidth of each underwater acoustic channel is set to 1 MHz. The oceanic noise is set to 10^{-8} W. The maximum transmission power of UWS is set to

1.5 W. The processing capability of UWS is set to 1 Gbps. The maximum processing capability of USV is set to 15 Gbps. The number of CPU cycles for processing one bit of data is set to 1000.

We evaluate the performance of the proposed algorithm with the benchmark algorithms. In random matching algorithm (RMA): the acoustic channels are randomly allocated to UWSs. In random offloading algorithm (ROA): the data is randomly offloaded to USV for processing. In random allocation computation resource algorithm (RACA): the computation resource of USV is randomly allocated to UWSs for processing data.

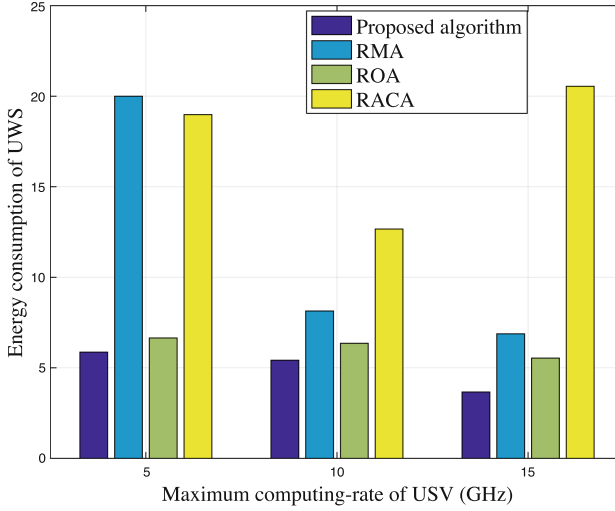


Fig. 2. Energy consumption of UWS with the changes of maximum computing-rate of USV.

Figure 2 shows the energy consumption of UWS with the changes of maximum computing-rate of USV, compared with the benchmark schemes. It can be seen that the proposed algorithm can obtain the low energy consumption for UWS. The reason is that the proposed algorithm jointly considers the acoustic channel allocation, data offloading, and the computation resource allocation of USV to improve the offloading efficiency.

Figure 3 depicts the energy consumption of UWS with the changes of data size. It can be obtained that the energy consumption of each UWS is increasing with the data size. The reason is that the local computation consumption and the offloading consumption will increase when the data size of each UWS is increased, which leads to the increase of energy consumption.

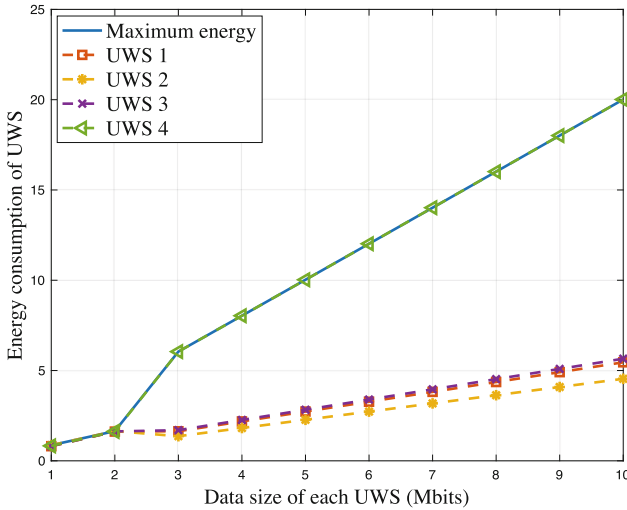


Fig. 3. Energy consumption of UWS with the changes of data size.

5 Conclusion

In this paper, we have investigated an energy efficient computation offloading scheme in maritime communication networks. In particular, we have proposed an optimization problem to minimize the total energy consumption of UWS by jointly optimizing the allocation of underwater acoustic channels, the offloading data, and the computation resource allocation of USV. To address the non-convexity of the joint optimization problem, we have designed a layered structure to derive the optimal solution. Simulation results have demonstrated the effectiveness and efficiency of our proposed algorithms. For the future work, we will plan to investigate the secrecy computation offloading for UWSs in maritime communication networks.

References

1. Qiu, T., Zhao, Z., Zhang, T., Chen, C., Chen, C.L.P.: Underwater internet of things in smart ocean: system architecture and open issues. *IEEE Trans. Industr. Inf.* **16**(7), 4297–4307 (2020)
2. Wei, T., Feng, W., Chen, Y., Wang, C.-X., Ge, N., Lu, J.: Hybrid satellite-terrestrial communication networks for the maritime internet of things: key technologies, opportunities, and challenges. *IEEE Internet Things J.* **8**(11), 8910–8934 (2021)
3. Arienzo, L.: Green RF/FSO communications in cognitive relay-based space information networks for maritime surveillance. *IEEE Trans. Cogn. Commun. Netw.* **5**(4), 1182–1193 (2019)
4. Zeng, H., Hou, Y.T., Shi, Y., Lou, W., Kompella, S., Midkiff, S.F.: A distributed scheduling algorithm for underwater acoustic networks with large propagation delays. *IEEE Trans. Commun.* **65**(3), 1131–1145 (2017)

5. Zhang, L., Liu, T., Motani, M.: Optimal multicasting strategies in underwater acoustic networks. *IEEE Trans. Mob. Comput.* **20**(2), 678–690 (2021)
6. Huang, X., Wu, K., Jiang, M., Huang, L., Xu, J.: Distributed resource allocation for general energy efficiency maximization in offshore maritime device-to-device communication. *IEEE Wirel. Commun. Lett.* **10**(6), 1344–1348 (2021)
7. Fang, F., Cheng, J., Ding, Z.: Joint energy efficient subchannel and power optimization for a downlink NOMA heterogeneous network. *IEEE Trans. Veh. Technol.* **68**(2), 1351–1364 (2019)
8. Yildiz, I.U., Gungor, V.C., Tavli, B.: Packet size optimization for lifetime maximization in underwater acoustic sensor networks. *IEEE Trans. Industr. Inf.* **15**(2), 719–729 (2019)
9. Song, Y.: Underwater acoustic sensor networks with cost efficiency for internet of underwater things. *IEEE Trans. Industr. Electron.* **68**(2), 1707–1716 (2021)
10. Su, X., Meng, L., Huang, J.: Intelligent maritime networking with edge services and computing capability. *IEEE Trans. Veh. Technol.* **69**(11), 13606–13620 (2020)
11. Fu, S., et al.: Energy-efficient UAV enabled data collection via wireless charging: a reinforcement learning approach. *IEEE Internet Things J.* **8**(2), 10209–10219 (2021)
12. Wu, Y., Ni, K., Zhang, C., Qian, L.P., Tsang, D.H.K.: NOMA-assisted multi-access mobile edge computing: a joint optimization of computation offloading and time allocation. *IEEE Trans. Veh. Technol.* **67**(12), 12244–12258 (2018)
13. Dai, M., Su, Z., Xu, Q., Zhang, N.: Vehicle assisted computing offloading for unmanned aerial vehicles in smart city. *IEEE Trans. Intell. Transp. Syst.* **22**(3), 1932–1944 (2021)
14. Yang, T., Feng, H., Yang, C., Wang, Y., Dong, J., Xia, M.: Multivessel computation offloading in maritime mobile edge computing network. *IEEE Internet Things J.* **6**(3), 4063–4073 (2019)
15. Ma, R., Wang, R., Liu, G., Meng, W., Liu, X.: UAV-aided cooperative data collection scheme for ocean monitoring networks. *IEEE Internet Things J.* **8**(17), 13222–13236 (2021)
16. Yang, T., et al.: Two-stage offloading optimization for energy-latency tradeoff with mobile edge computing in maritime internet of things. *IEEE Internet Things J.* **7**(7), 5954–5963 (2020)
17. Stojanovic, M.: On the relationship between capacity and distance in an underwater acoustic communication channel. In: *Proceedings of ACM International Workshop Underwater Networking*, pp. 41–47 (2006)
18. Qian, L., Wu, Y., Yu, N., Jiang, F., Zhou, H., Quek, T.Q.S.: Learning driven NOMA assisted vehicular edge computing via underlay spectrum sharing. *IEEE Trans. Veh. Technol.* **70**(1), 977–992 (2021)
19. Wu, Y., Song, Y., Wang, T., Qian, L., Quek, T.Q.S.: Non-orthogonal multiple access assisted federated learning via wireless power transfer: a cost-efficient approach. *IEEE Trans. Commun.* **70**(4), 2853–2869 (2022)



Game Theoretic Analysis of Resource Allocation in Multi-tiered Networks

Li Zhang^(✉), Mohit Hota, and Sanjiv Kapoor

Illinois Institute of Technology, Chicago, IL 60616, USA
{lzhang115,mhota}@hawk.iit.edu, Kapoor@iit.edu

Abstract. We study resource allocation in wireless multi-tiered OFDMA cellular systems, like the current 5G networks. We compare coordinated versus distributed methods of assignment of resource blocks to maximize signal strength and throughput to ensure quality of service. The distributed assignment, where every wireless client optimizes its utility, is modeled by a strategic game with the cellular users as players. We study the existence of Nash equilibrium in this game and provide bounds for the inefficiency of this strategic game via analysis of the Price of Anarchy (PoA). Our social objective is chosen to provide fair throughput by measuring the minimum data rate. We show that approximating the social optimum, i.e. maximizing the minimum data rate, to within any given factor is NP-Hard and provide an algorithm that is illustrated to perform well in practice.

Keywords: Multi-tiered networks · Price of anarchy · Nash equilibrium

1 Introduction

Next generation cellular networks, that include 5G networks, provide enhanced performance with high speeds for downlinks and uplinks via a multi-tiered architecture. The architecture utilizes macrocells, licensed small cells and device-to-device networks [3] integrating use of multiple radio access technologies amongst others.

In the context of a multi-tiered architecture, with increased number of users and connected devices [2], distributed methods that pair user equipment to base stations are becoming increasingly important keeping in view enhanced quality of service. In this paper, the resource allocation problem is investigated in the context of a multi-tier *orthogonal frequency division multiple access* (OFDMA) network. User equipment connect to base stations utilizing resource blocks corresponding to frequency and time partitions. We consider the performance of distributed systems that pair cellular users to resource blocks in order to optimize their own data rate.

A number of distributed methods for the resource allocation method have been proposed including the use of matching theory, factor graphs and optimization using auctions [3]. The approach of distributed allocation presents itself

naturally as a game where users adopt strategies to maximize their benefit. Our goal in this paper is to analyze this game and the inefficiency that result from the strategic behavior of the players at Nash equilibrium. Our social objective is to maximize the minimum throughput, measured as a function of signal-to-interference-plus-noise ratio (SINR) that the players are able to achieve, so as to ensure a fair distributions of resources. We utilize the *price of anarchy* as a measure of the inefficiency that results from the strategic equilibrium choices in the game.

Prior work includes multiple papers that attempt to optimize resource allocation. A stable matching based algorithm is designed in [6] to allocate the resources to the uplink transmitters in a 5G cellular networks. Optimizing the power utilization has also been considered [7] especially in the context of device-to-device (D2D) networks, aimed at minimizing the transmission power of all the users in the network. The research does not account for multi-tiered architectures.

A resource allocation problem focused on D2D communications and formulated in the context of 5G Radio Access Networks has been considered in [9] and shown to be NP-hard. The optimization criterion is the sum of the capacities of all the communication links and a fair allocation problem has not been considered. The paper also proposed a many-to-one matching coalition sub-game to model the problem and determine stable solutions. Maximizing resource allocation in the context of Mobile Edge Computing (MEC) has been considered in [8].

Game theoretic modeling of resource allocation has also been considered in [4] where a Stackelberg game has been considered for power allocation in the context of soft frequency reuse (SFR), with an interference coordination method for decreasing inter-cell interference. Utilization of network slicing techniques [1] for virtualization and dynamically changing the resource allocation to optimize the quality of service and experience has also been considered.

1.1 Our Contributions

1. We formulate the resource allocation problem as a strategic game, the *Wireless Resource Game* with small cell users as players. Our work differs from previous work in considering mobile clients that are allowed to switch base stations.
2. We show that Nash equilibrium is not always possible and strategic play can result in an unstable system, when mobile clients are allowed to optimize their power levels and choice of base stations.
3. We provide bounds on the inefficiency of the strategic game as measured by the *price of anarchy* (PoA). The price of anarchy is dependent on the number of wireless clients and the ratio of the maximum to minimum signal strength. We illustrate the average behavior of the price of anarchy via a simulation that utilizes a geographically located system of small cell users and base stations.

4. We show that the optimum social objective problem is NP-Hard to approximate to within a factor of L where, theoretically, L can be arbitrary. We also show the performance of a simple swap based algorithm for determining a local optimum of the social objective function.

2 Network Model and Game Formulation

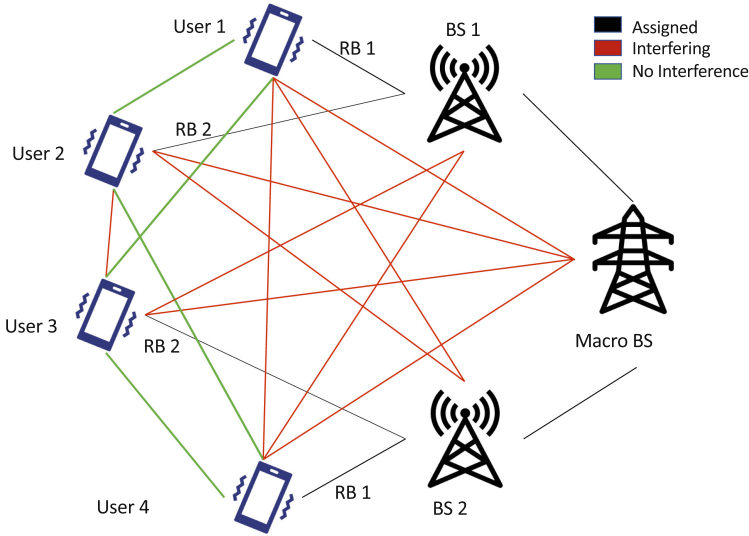


Fig. 1. Illustration of a multi-tiered cellular network

Our model for a 5G cellular wireless system includes a macro base station (MBS) and a set of small cell base stations (SBS) used by users' equipment (UEs). There is a set of fixed power levels and resource blocks (RBs) available to base stations and users. The objective of the resource allocation problem is to allocate resource blocks to small cell base station users to optimize signal quality (SINR), and hence the throughput capacity, while accounting for interference from other users and base stations.

In a distributed allocation system, each user equipment is allowed to choose a SBS for communication. This gives rise to local optimization (and selfish behavior) by the users which we model via a strategic game. Users select a resource block each, from among those available, in an effort to maximize their own spectral efficiency measured as a function of SINR, accounting for the possible interference from other users and MBS's. We study properties and characteristics of Nash equilibrium in these games, and compare the efficiency with a socially optimum fair assignment of users.

In our model, we let K be the set of small cell user equipments, where each small cell user is a transmitter belonging to set K^T and receiver belonging to K^R . The set of BSs is denoted by B . Also, small cell users are able to select the transmission power from a finite set of power levels and also choose RBs $R = \{1, 2, \dots, |R|\}$. The possible power over the RB for each transmitter k can be marked as the vector $P_k = [P_k^1, P_k^2, \dots, P_k^{|R|}]^T$ where $P_k^r = \{p_{k_1}^r, p_{k_2}^r \dots p_{k_1}^r\} \geq 0$ represents the power level of transmitter k over RB r which was selected from the set of power levels. And $p_k^r = 0$ if the RB r is not allocated to the transmitter. Based on the description of SBS utilization in [5] we assume that each RB in each SBS can be occupied by only one user.

As an example, consider the physical model illustrated in Fig. 1. The network consists of two SBSs (BS1 and BS2) and four small cell users. The available frequencies are divided over time periods into resource blocks. Users 1 and 2 select $RB1$ and $RB2$ on $BS1$ and users 3 and 4 select $RB2$ and $RB1$ on $BS2$ respectively. Thus, User 2 and 3 interfere with each other since they use the same resource block on two different base stations, as do users 1 and 4. That is, there is no interference between BSs or users when different RBs are being used, and interference arises due to simultaneous use of resource blocks between users or BSs.

The terms used in this paper are described in Table 1.

The SINR, denoted by $\gamma_{b,k}^r$, and the utility $u_{b,k}^r$ of a small cell user k occupying BS b and RB r are respectively defined as:

$$\gamma_{b,k}^r = \frac{I_{b,k}^r}{G_k^r - I_{b,k}^r} \text{ and } u_{b,k}^r = \log_2(1 + \gamma_{b,k}^r) \quad (1)$$

A *Wireless Resource Game* is defined as $\mathcal{W}(K, B, R, U)$, where K is set of users, R is set of RBs, B is set of BSs, and U is the set of user utilities for a given allocation of users to resource blocks and base stations. We assume that the user utilizes her maximum power level, termed p_k^r , since there is no penalty incorporated for power usage in our current model. The strategies of the players are given by strategy profile $S = (S_k), \forall k$ where S_k is the strategy set of user k , given by the choice of the BS b and RB r , i.e., $S_k = \{(r_i, b_j)_{i \in R, j \in B}\}$. The utility function for each user k , $U_k : S \rightarrow \mathcal{R}^+$ represents the utility $u_{b,k}^r$ of the user when utilizing RB r on BS b , given the strategies of the other players. The user selects a BS and RB to maximize her utility based on the strategies of other small cell users. This gives rise to the following optimization problem for small cell user $k \in K$:

Optimization Problem 1:

$$\begin{aligned} \max \quad & \sum_r \sum_b x_{k,b}^r \log_2(1 + \gamma_{b,k}^r) \\ \text{s.t.} \quad & \forall k : \sum_r \sum_b x_{k,b}^r \leq 1 \\ & \forall k, b, r : x_{k,b}^r \in \{0, 1\} \end{aligned} \quad (2)$$

Table 1. Table of parameters

Symb	Description
σ^2	noise
$g_{k,b}^r$	Channel gain between transmitter k and receiver b over RB r
p_k^r	Transmission power level of user k over RB r
p_b^r	Transmission power level of BS b over RB r
$\Gamma_{b,k}^r$	Signal strength at receiver k over RB r from transmitter b
$\Gamma_{M,k}^r$	Signal strength of interference at receiver k over RB r from MBS M
$\Gamma_{k',k}^r$	Signal strength of interference at receiver k over RB r from transmitter k' (on a separate BS)
G_k^r	Cumulative (sum of desired signal strength, interference, and noise) signal strength of received by user k over RB r
Γ_{min}	Lowest value of the signal strength of all users
Γ_{max}	Highest value of the signal strength of all users
R	Finite set of RBs $R = \{1, 2, \dots, R \}$
B	Finite set of BSs $B = \{1, 2, \dots, B \}$
b_k	BS that user k occupies
K_b	Users K that occupy BS b
Γ_{max}^r	Maximum interference in RB r
C_b^r	Capacity on BS b in RB r

$$\Gamma_{b,k}^r = g_{b,k}^r p_b^r; \quad \Gamma_{M,k}^r = g_{M,k}^r p_M^r$$

$$G_k^r = \sum_{k' \neq k} \Gamma_{k',k}^r + \sum_{b_{k'} \neq b} \Gamma_{b_{k'},k}^r + \Gamma_{b,k}^r + \Gamma_{M,k}^r + \sigma^2$$

$$\Gamma_{min} = \min_{k \in K^T, b \in B, r \in R, s \in S} \Gamma_{b,k}^r; \quad \Gamma_{max} = \max_{k \in K^T, b \in B, r \in R, s \in S} \Gamma_{b,k}^r$$

where $x_{k,b}^r = 1$ if small cell user k utilizes resource x block r on the SBS b , while $x_{k,b}^r = 0$ otherwise. In this non-cooperative game for small cell users, the stable solution is a Nash equilibrium, NE (if it exists). A strategy profile S^{NE} is a Nash equilibrium strategy if:

$$U_k(S_k^{NE}, S_{-k}^{NE}) \geq U_k(S'_k, S_{-k}^{NE})$$

where S_k^{NE} is the resource utilization strategy for player k , S'_k is any other strategies besides Nash for user k and S_{-k}^{NE} is the strategy for the rest of the players at Nash equilibrium, $U_k = u_{b,k}^r$ when $x_{k,b}^r$ denotes users k 's utility. Additional

constraints on bandwidth capacity of the base stations and interference can be expressed as:

$$\forall r, b : \sum_{k \in K_b} x_{k,b}^r \log_2(1 + \gamma_{b,k}^r) \leq C_b^r$$

$$\sum_{k' \in K^T; k' \neq k} g_{k',k}^r p_{k'}^r + \sum_{b' \neq b} g_{b',k}^r p_{b'}^r + g_{M,k}^r P_M^r \leq I_{max}^r \tag{3}$$

We use Shannon’s formula, $C = B * \log_2(1 + \frac{S}{N})$, to relate SINR and link capacity.

We analyze the solution obtained from the Nash equilibrium via the social objective function that measures the minimum utility achieved by the users, i.e., $U^{soc} = \min_k \sum_b \sum_r x_{k,b}^r \log_2(1 + \gamma_{b,k}^r)$, where $x_{k,b}^r$ is the indicator function described above and U^{soc} is the social utility. This is termed the *fair* metric. The social optimum can be found using the following problem:

Optimization Problem 2: $\max \min_k \sum_b \sum_r x_{k,b}^r \log_2(1 + \gamma_{b,k}^r)$ with additional constraints as in Optimization Problem 1 and in (3)

3 Nash Equilibrium and Price of Anarchy

In this section we consider existence of Nash equilibrium. Unfortunately, the game can be unstable as we show below an instance where Nash equilibrium may not exist.

Claim. For the Wireless Resource Game, $\mathcal{W}(K, B, R, U)$ Nash Equilibrium does not always exist.

Proof. An example where Nash equilibrium does not exist is created using a random sampling Monte Carlo method. The example has 3 users and 2 RBs with only 3 BSs. We assume that there is only one MBS. The following values of interference signal strengths provide an example where no strategy is stable.

$$\begin{bmatrix} k_1(r_1) & k_1(r_2) & k_2(r_1) & k_2(r_2) & k_3(r_1) & k_3(r_2) \\ k_1 & & 22.92 & 28.24 & 21.94 & 16.50 \\ k_2 & 22.92 & 28.24 & & 22.39 & 26.77 \\ k_3 & 21.94 & 16.50 & 22.39 & 26.77 & \end{bmatrix}.$$

The value of noise is 0.30698 mW and assuming there is no interference from MBS.

In this example, there is no configuration that is a Nash equilibrium. For instance, if user k_1 occupies b_1 and r_1 , k_2 occupies b_1 and r_2 , k_3 occupies b_2 and r_1 , then k_1 is willing to move to b_2 with r_2 . Similarly, we can check that for any

configuration, at least one user is willing to change its strategy in any of the configuration. \square

Though the Nash equilibrium does not always exist in the general Wireless Resource Game, it does exist for special game classes as illustrated below:

Uniform Interference Property: In this case we assume that each client has the same signal strength for all RBs and BS, i.e. $\Gamma_{k,b}^r = a_k$ for all k, b, r , the same interference signal strength from other users, i.e. $\Gamma_{k',k}^r = b_k$ for all k, k', r and the same interference signal strength from any BS, i.e. $\Gamma_{b',k}^r = c_k$ for all b', k, r . Also assume that the interference signal strength from the MBS is the same for all users, i.e. for all k, r , $\Gamma_{M,k}^r = a_M$.

Lemma 1. *A Nash equilibrium always exists for a Wireless Resource Game with Uniform interference property.*

Proof. Suppose a specific user is fine with a particular signal strength, and she is willing to move to any RB and BS as long as such condition has been satisfied. In this situation, for each user k , its signal strength is the same among all RBs and BSs $\Gamma_{k,b}^r = a_k$. We can claim that the user will switch to another RB and BS if and only if the load on other RBs is less than that of the user is currently in. In this game, $\frac{a_k}{g_k^{r_1} - a_k} \leq \frac{a_k}{g_k^{r_2} - a_k} \implies g_k^{r_1} \geq g_k^{r_2}$. Since the interference from all the other users and BSs are the same, $g_k^{r_1} = (n_{r_1} - 1) * (b_k + c_k) + a_m$ where n_{r_1} is the number of users on RB r_1 . If user k is willing to move to r_2 from r_1 , $n_{r_1} - 1 > n_{r_2}$. So let n as the list of the number of users on each RB in order from highest to lowest, i.e. $n = (n_1, n_2, \dots, n_r)$ where n_i means i_{th} highest number of users among all RBs, we can find that after the users switch RBs the new list $n' = (n'_1, n'_2, \dots, n'_r)$ is in a decreasing lexicographic order, which implies convergence to Nash equilibrium. \square

3.1 Inefficiency of Equilibrium

We provide a bound on the inefficiency of the *Wireless Resource Game* using the *Price of Anarchy*. The PoA is defined as follows:

$$PoA = \inf_G \frac{\min_{NE(G)} \min_k U_k(S_k^{NE(G)}, S_{-k}^{NE(G)})}{\min_k U_k(S_k^*(G), S_{-k}^*(G))}$$

where G is a game with strategy profile S , $S_k^{NE(G)}$ is a Nash strategy for user k in game G , $S_k^*(G)$ is the social optimal strategy for user k in game G , $U_k(S)$ the utility of user k under the strategy profile S . We are using the fair metric, the social utility being defined as the minimum utility among all users.

The Theorem below illustrates the dependency on the ratio of the signal strength of the users and indicates that it should be kept bounded. It is also dependent on the number of users in the model.

Theorem 1. 1. The PoA of the Wireless Resource Game using the fair metric and without capacity and interference constraints (without Eq. 3) is lower bounded as

$$PoA \geq \frac{1}{2|K|(\Gamma^m)^2(1 + \log_2(\Gamma^m))} \text{ where } \Gamma^m = \frac{\Gamma_{max}}{\Gamma_{min}}.$$

2. The PoA of the Fair-Wireless Resource Game with capacity and interference constraints (Eq. 3) is lower bounded as

$$PoA \geq \frac{1}{C_{b_{max}}} \log_2 \left(1 + \frac{\Gamma_{max}}{\Gamma^m I_{max}} \right)$$

where $C_{b_{max}} = \max_{r \in R} C_b^r$ and $I_{max} = \max_{r \in R} I_{max}^r$.

Proof. We first consider the case where interference and capacity constraints do not exist. Let r_k^{NE} denote the RB user k uses at a Nash equilibrium, r_k^* denote the RB at a social optimum, and G_k^r and $G_k^{r'}$ denote the total signal strength received by user k in RB r and r' , respectively, which is the sum of interference signals plus the signal strength from its own BS. Then, for any RB r , we have

$$\begin{aligned} \frac{\Gamma_{b,k}^{r_{NE}}}{G_k^{r_{NE}} - \Gamma_{b,k}^{r_{NE}}} &\geq \frac{\Gamma_{b',k}^{r_{NE}}}{G_k^{r_{NE}} - \Gamma_{b',k}^{r_{NE}}} \implies G_k^{r_{NE}} - \Gamma_{b,k}^{r_{NE}} \leq \Gamma^m (G_k^{r_{NE}} - \Gamma_{b',k}^{r_{NE}}) \\ G_k^{r_{NE}} - \Gamma_{b,k}^{r_{NE}} &\leq \Gamma^m (\max(\sum_{k' \neq k} \Gamma_{b_{k'},k}^{r'} + \Gamma_{k',k}^{r'} + \Gamma_{M,k}^r) + \sigma^2) \end{aligned}$$

where $\Gamma^m = \Gamma_{max}/\Gamma_{min}$

For the user with the least signal strength using the RB r_{NE} , the lower and upper bounds on SINR (Eq. 1) are expressed as

$$\begin{aligned} \gamma_k^{NE} &= \frac{\Gamma_{min}}{\Gamma^m (\max(\sum_{k' \neq k} \Gamma_{b_{k'},k}^{r'} + \Gamma_{k',k}^{r'} + \Gamma_{M,k}^r) + \sigma^2)} \\ \text{and } \gamma_k^* &= \frac{\Gamma_{max}}{\Gamma_{M,k}^r + \sigma^2} \leq \frac{\Gamma_{max}}{\Gamma_{min} + \sigma^2} \end{aligned}$$

Thus, the PoA is given by

$$\begin{aligned}
 PoA &\geq \frac{\log_2 \left(1 + \frac{\Gamma_{\min}}{\Gamma^m \left(\max_{k' \neq k} \left(\sum \Gamma_{b'_{k',k}}^{r'} + \Gamma_{k',k}^{r'} + \Gamma_{M,k}^r \right) + \sigma^2 \right)} \right)}{\log_2 \left(1 + \frac{\Gamma_{\max}}{\Gamma_{\min} + \sigma^2} \right)} \\
 &\geq \frac{\frac{\Gamma_{\min}}{\Gamma^m \left(\max_{k' \neq k} \left(\sum \Gamma_{b'_{k',k}}^{r'} + \Gamma_{k',k}^{r'} + \Gamma_{M,k}^r \right) + \sigma^2 \right)}}{\log_2 \left(\frac{2\Gamma_{\max}}{\Gamma_{\min} + \sigma^2} \right)} \\
 &\geq \frac{\frac{\Gamma_{\min}}{\Gamma^m (2(|K|-1)\Gamma_{\max} + \Gamma_{\max} + \sigma^2)}}{\log_2 \left(\frac{2\Gamma_{\max}}{\Gamma_{\min} + \sigma^2} \right)} \geq \frac{\frac{\Gamma_{\min}^2}{\Gamma_{\max}((2|K|-1)\Gamma_{\max} + \sigma^2)}}{1 + \log_2(\Gamma^m)} \\
 &\geq \frac{\frac{\Gamma_{\min}^2}{\Gamma_{\max}((2|K|-1)\Gamma_{\max} + \Gamma_{\max})}}{1 + \log_2(\Gamma^m)} \geq \frac{\frac{\Gamma_{\min}^2}{2|K|\Gamma_{\max}^2}}{1 + \log_2(\Gamma^m)} \\
 &\geq \frac{\frac{1}{2|K|\Gamma^m}}{1 + \log_2(\Gamma^m)} \geq \frac{1}{2|K|\Gamma^m(1 + \log_2(\Gamma^m))}
 \end{aligned}$$

The second inequality holds because the lower bound of γ_k^{NE} can be very small, also we assume that the noise σ is negligible.

PoA with Capacity and Interference Constraints: We start with the previous analysis of $G_k^{r_{NE}}$. The lower bound for user k in r_{NE} is when the user has the least signal strength while experiencing the largest interference I_{max} :

$$\gamma_k^{NE} = \frac{\Gamma_{\min}}{\Gamma^m(I_{max} - \Gamma_{\min} + \sigma^2)}$$

The upper bound for this user in RB r^* can be calculated based on the capacity constraint: $\gamma_k^* \leq 2^{C_{b_{max}}} - 1$ where $C_{b_{max}} = \max_{r \in R} C_b^r$. The result follows.

$$\begin{aligned}
 PoA &\geq \frac{1}{C_{b_{max}}} \log_2 \left(1 + \frac{\Gamma_{\min}}{\Gamma^m(I_{max} - \Gamma_{\min} + \sigma^2)} \right) \\
 &\geq \frac{1}{C_{b_{max}}} \log_2 \left(1 + \frac{\Gamma_{\min}}{\Gamma^m(I_{max} + \sigma^2)} \right) \\
 &\geq \frac{1}{C_{b_{max}}} \log_2 \left(1 + \frac{\Gamma_{\min}^2}{\Gamma_{max}(I_{max} + \sigma^2)} \right)
 \end{aligned}$$

□

We now show that the lower bound presented in the above theorem can be arbitrarily small. This requires a wide variation in signal strength, which in practice should be kept bounded.

Claim. There exists an instance of the the Wireless Resource Game, when using the fair metric, where $PoA = \frac{1}{\log(1+M)}$ where M is arbitrarily large.

Proof. We provide an example with 2 users and 2 RBs. We ignore interference from all BSs including MBS, and assign signal values to achieve an arbitrarily small PoA. Consider values $\Gamma_{11}^1 = \Gamma_{21}^2 = M$, and $\Gamma_{21}^1 = 1, \Gamma_{11}^2 = M$. We assume that interference from noise, σ^2 is 1. The optimum assignment is when user 1 is assigned to RB 1 and user 2 to RB 2; the minimum utility and thus the social objective function is $\log(1 + \frac{M}{\sigma^2})$. It can be verified that a Nash equilibrium assignment exists when user 1 is assigned to RB 2 and user 2 is assigned to RB 1. The social objective function at this Nash equilibrium evaluates to $\log(1 + \frac{1}{\sigma^2})$. The result follows. \square

4 Determining Social Optimum

In this section we show that determining the social optimum, even approximately, is NP-Hard. However, we also present an efficient heuristic to discover social optimum.

Theorem 2. *Determining an approximate social optimum in the wireless resource allocation problem using the max-min measure is NP-Hard.*

Proof. We reduce from the known NP-Hard problem of the *chromatic number problem* where we must determine if a graph G is q -colorable. From a given instance $I_c = (G(V, E), q)$ of the chromatic number problem, we create an instance of our 5G wireless game $I_w = (K, B, R)$ - For each node v in G , we create a user k_v in K . There are q RBs and $|B| = n * q$ base stations (n is the number of nodes in the graph), and each user k faces an interference of I_B from every base station she does not occupy. If an edge exists between two nodes u and v , the corresponding users k_u and k_v heavily interfere with each other if they share the same RB; where we define *heavy interference* to be the case when the interference signal is no less than the given value $I_B|B|(L + 2)$. If two nodes are not neighbors, the corresponding users interfere with each other with value δ , where δ is a small positive value. The signal strength of each user is $g_{k,b}^r p_k^r = 1, \forall r, b$. Therefore, a user k facing no heavy interference obtains a utility bound

$$\gamma_{k,b}^r \geq \frac{1}{\sigma^2 + \delta|B| + I_B|B|},$$

and a user facing heavy interference from at least one other user obtains a utility bound

$$\gamma_{k,b}^r \leq \frac{1}{\sigma^2 + I_B|B|(L + 2) + I_B(|B| - 1) + (n - 2)\delta}.$$

Given a social optimum, S_w in I_w , we show an equivalent solution for q -coloring I_c . If the social utility is $\max \min_k \gamma_{k,b}^r \geq \frac{1}{\eta + \delta|B| + I_B|B|}$, no two users that heavily interfere with each other can occupy the same resource block, since their utilities

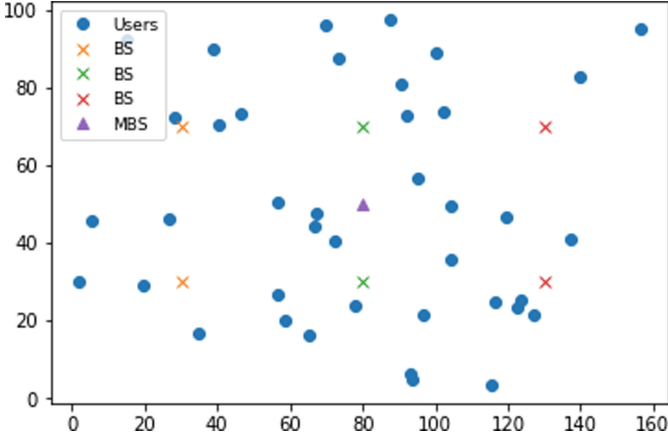


Fig. 2. Geometric distribution of Users and Base stations

would then be lesser. Then, for each RB r , no two occupying users will have their corresponding nodes in I_c as neighbors, i.e., the corresponding nodes of users in each RB r will be an independent set I_r . The converse can be proved similarly, implying that the problem is NP-Hard.

We now show that approximating the problem of finding the socially optimal solution is also NP-Hard. In I_w , consider the utility of a user k_v using RB r on BS b with no high-value interference from another user k_u (corresponding nodes in I_c are not neighbors). Assuming the noise σ^2 to be negligible, the bound on k_v 's utility is then $\gamma_c \geq \frac{1}{\sigma^2 + \delta|B| + I_B|B|}$ since the same resource block may be occupied on every possible BS. Similarly, if a user k_v faces heavy interference from no less than one other user (corresponding nodes in I_c are neighbors), her utility will be bounded as $\gamma_{nc} \leq \frac{1}{\sigma^2 + I_B|B|(L+2)}$. Noting that L , n and $|B|$ are sufficiently large constants, the utilities u_c and u_{nc} corresponding to γ_c and γ_{nc} respectively are $u_c \geq \log(1 + \gamma_c) = \log\left(1 + \frac{1}{\delta|B| + I_B|B|}\right) \approx \frac{1}{\delta|B| + I_B|B|}$, and similarly $u_{nc} \leq \log(1 + \gamma_{nc}) \approx \frac{1}{I_B|B|(L+2)}$. Let $\delta = \frac{I_B}{L+1}$, the ratio between the two utility bounds is then

$$\frac{u_c}{u_{nc}} \geq \frac{I_B|B|(L+2)}{\left(1 + \frac{1}{L+1}\right)I_B|B|} \geq \frac{L+2}{1 + \frac{1}{L+1}} > L.$$

This gap can be used to show that finding an approximation better than factor $\frac{1}{L}$ is NP-Hard. □

4.1 Exchange-Method for Social Optimum

We now present Algorithm 1 which iteratively improves the social objective function. At each step, all users are considered in a round-robin order (line 6) and

it is determined if the user can improve the social objective function by changing his own assignment (line 10) or by switching his assignment with another user's assignment (line 14) . The algorithm terminates when no improvements are possible.

Algorithm 1: Exchange-Method

```

1  $u_{old} = -1, u_{new} = 0$ 
2 Assign all users randomly to BSs and RBs.
3 Let list  $Q$  contain all unoccupied RB and BS pairs.
4 while  $u_{new} \neq u_{old}$  do
5      $u_{old} = u_{new}$ 
6     for  $j \in 1 \dots |K| - 1$  do
7          $\forall (r, b) \in Q$ , let  $A_{(r,b)}$  be the assignment when user  $k_j$  is assigned to
             $(r, b)$ ; calculate minimum utility over all users:
             $u(A_{(r,b)}) = \min_k u_{k,b}^r$ .
8          $(r^*, b^*) = \arg \max_{(r,b) \in Q} u(A(r, b))$ .
9         Let  $u'_{min}$  be the minimum utility achieved after  $k_j$  is assigned
             $(r^*, b^*)$ .
10        if  $u'_{min} > u_{new}$  then
11            Let  $k_j$  picks the strategy  $(r^*, b^*)$  such that  $u_{r,b}^{k_j} = u'_{min}$ .
12             $u_{new} = u'_{min}$ 
13        else
14            For all pairs  $(k, k_j)$  where  $k \neq k_j$ : check if switching  $k$  with  $k_j$ 
                improves  $u_{min}$ 
15             $u_{new} = u_{min}$ 

```

Approximation Ratio of the Exchange-Method: Note that the approximation ratio can be arbitrarily large - Consider an example with 3 users and 3 RBs,

where δ is an arbitrarily small number, with signal strength matrix $\Gamma = \begin{bmatrix} 1 & \delta & \delta \\ \delta & 1 & \delta \\ \delta & \delta & 1 \end{bmatrix}$

where rows are users and columns are RBs. If we start with an assignment where k_2 is using r_1 , k_3 is using r_2 and k_1 is using r_3 , then each user has a signal strength of δ , i.e., user with minimum utility gets δ . The algorithm terminates with this solution. However, the optimal social solution would be the assignment with k_1 using r_1 , k_2 using r_2 and k_3 using r_3 , leading to the ratio of best-to-worst solution of the exchange method as $\frac{1}{\delta}$.

Complexity of the Exchange-Method: We note that the minimum utility has an upper bound, given by $\frac{\Gamma_{max}}{\sigma^2}$; and therefore, the exchange-method algorithm always terminates. Let T be the maximum number of strategy changes or user pairs switching. The inner *for* loop runs $|K|$ times and line 6 takes time proportional to $O(|R||B||K|)$. Line 13 takes time proportional to $O(K^2)$, implying a total algorithm running time of $O(T|K|^2|R||B|)$.

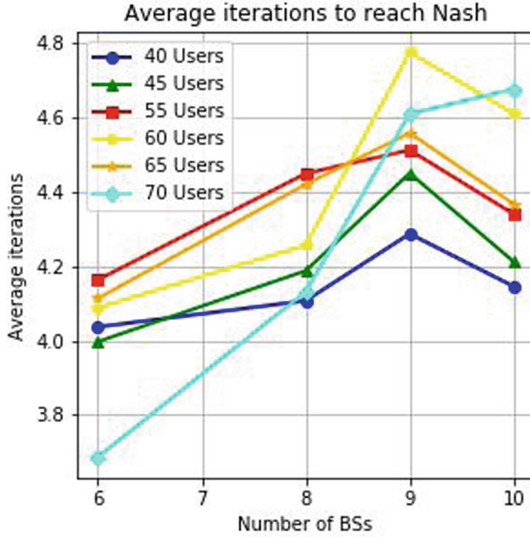


Fig. 3. Effect of number of Users and BSs on number of iterations on finding Nash equilibrium

5 Simulations

In this section, we present simulations to study Nash equilibrium and price of anarchy in a realistic geographic setting. We provide experimental results on the best response method, examining the speed of convergence and the success rate of achieving Nash equilibrium. We also experimentally analyze the complexity of Algorithm 1 in terms of number of exchanges performed before termination.

The 5G wireless communication model consists of 40 to 65 users, 12 RBs and 6 to 10 BSs as is typical for these models, the reason we don't set number of RBs to a big number is to avoid the situation that most RBs are occupied by only one or two users with little interference. Users are located within a rectangular space of size $100 * 160$ m, the BSs are located on a grid contained within that space (illustrated in Fig. 2), and the MBS is located at the center of the map. The channel gain ranges from -15 dB to -31 dB and power ranges from 26 dBm to 50 dBm. These data ranges are based on a realistic 5G network. For our simulations, we pick random numbers within the specified ranges for each run.

5.1 Simulations to Analyze Nash Equilibrium and Price of Anarchy

The Nash equilibrium is determined via the best-response method as follows: Users are considered in a round-robin order to determine if they will switch out of the currently assigned resource block and the base station to an alternate resource block and base station pair, the pair being chosen in order to maximize the user's utility given the strategies of the other players.

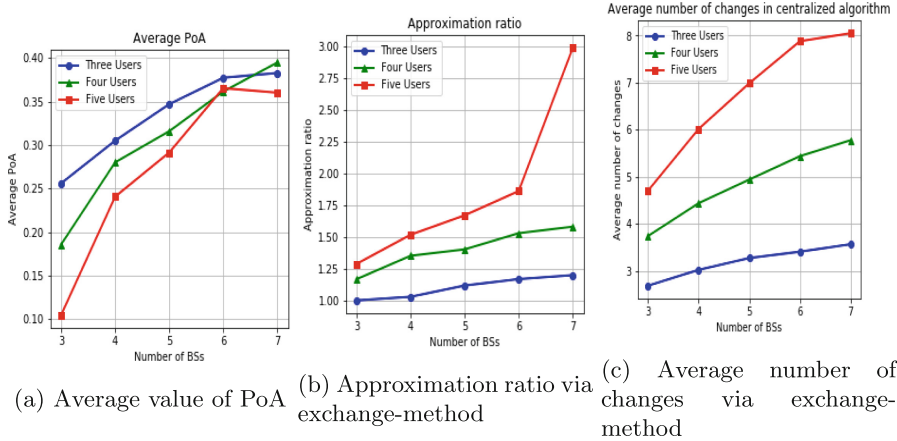


Fig. 4. Effect of number of Users and BSs on inverse of the PoA and approximation ratio using exchange-method

The number of users in our simulations range from 40 to 65 in intervals of size 5, and the number of BSs range between 6 and 10. Figure 3 shows the number of iterations required to reach Nash equilibrium. When the algorithm converges to a Nash equilibrium, as it does in 99% of cases in our experiments, we observe that the number of iterations is at most 5, and generally increases with an increase in the number of users.

For PoA, we analyze for only a small number of users (ranging from 3 to 5) and base stations (ranging from 3 to 7) since computing the optimum solution using the max-min social objective is NP-hard. Figure 4a illustrates the average value of the PoA value achieved for with different number of users and BSs while the number of RBs are set to 2. We observe that the PoA (for the max-min social objective measure) increases from 0.105 to 0.4 as the number of BSs increases. This indicates that increase in number of BSs improves the value of the PoA, as expected. In this setting, we also consider the exchange method and compare the results with the actual social optimum, which we compute via a complete case analysis. Figure 4b shows the median approximation ratio defined as the ratio of the optimum social objective value to the social objective value obtained using Algorithm 1. We observe that this ratio is bounded (in fact at most 3 for the cases considered); clearly as the number of users and the number of BSs increase, the approximation ratio increases, i.e., the performance of the algorithm becomes worse.

Figure 4c illustrates the average number of the assignment changes being made during the implementation of the exchange-method. The number of changes is the sum of the number of times a user changes its strategy and the number of times two users swap assignments. We note that the increase in the number of changes is almost linear w.r.t the number of users and BSs; the number of average changes are 2.69, 3.74 and 4.71 as the number of users increase from 3 to 5 when there are 3 BSs.

6 Conclusions

We have studied 5G communication networks and developed a game model comprising mobile clients in a distributed setting. We showed that finding the social optimal using max-min metric is NP-Hard and provided an exchange-method heuristic. We also provided a best response method to determine Nash equilibrium. Simulation results show the effectiveness of our analyses and algorithms.

References

1. Dighriri, M., et al.: Resource allocation scheme in 5G network slices. In: 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 275–280. IEEE (2018)
2. Gallagher, J.C., DeVine, M.E.: Fifth-generation (5G) telecommunications technologies: issues for congress. *Congr. Res. Serv.* **1(30)**, 1–39 (2019)
3. Hasan, M., Hossain, E.: Distributed resource allocation in 5G cellular networks. In: *Towards 5G: Applications, Requirements and Candidate Technologies*, pp. 245–269 (2015)
4. Huo, L., Jiang, D.: Stackelberg game-based energy-efficient resource allocation for 5G cellular networks. *Telecommun. Syst.* **72(3)**, 377–388 (2019)
5. Kwan, R., Leung, C., Zhang, J.: Multiuser scheduling on the downlink of an LTE cellular system. *Rec. Lett. Commun.* **2008** (2008)
6. Mathur, R.P., Pratap, A., Misra, R.: Distributed algorithm for resource allocation in uplink 5G networks. In: *Proceedings of the 7th ACM International Workshop on Mobility, Interference, and MiddleWare Management in HetNets*, pp. 1–6 (2017)
7. Omorinoye, A., et al.: On the resource allocation for D2D underlaying uplink cellular networks. In: *2019 26th International Conference on Telecommunications (ICT)*, pp. 108–112. IEEE (2019)
8. Pereira, R.S., et al.: RELIABLE: resource allocation mechanism for 5G network using mobile edge computing. *Sensors* **20(19)**, 5449 (2020)
9. Zhang, B., et al.: Resource allocation for 5G heterogeneous cloud radio access networks with D2D communication: a matching and coalition approach. *IEEE Trans. Veh. Technol.* **67(7)**, 5883–5894 (2018)



Energy-Efficient Multi-cell NOMA Design via Coalition Formation Game

Yanqing Xu¹, Donghong Cai²(✉), Shuai Wang¹, Zhicheng Dong³,
and Weixi Zhou⁴

¹ The School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China

xuyanqing@cuhk.edu.cn, shuaiwang@link.cuhk.edu.cn

² The College of Cyber Security, Jinan University, Guangzhou 510632, China
dhcai@jnu.edu.cn

³ College of Information Science and Technology, Tibet University, Lhasa, China
dongzc666@163.com

⁴ School of Computer Science, Sichuan Normal University, Chengdu, China
zhouweixi@sicnu.edu.cn

Abstract. In this work, we consider a downlink non-orthogonal multiple access (NOMA) network where multiple single-antenna users are served by multiple multi-antenna base stations (BSs). For practical considerations, we assume that only the imperfect channel state information (CSI) of each user is available at the BSs. Based on this model, the problem of joint user grouping and robust beamforming design is formulated to minimize the sum transmission power, and meanwhile, guarantee the quality of service requirements of users. Due to the integer variables of user grouping, coupling effects of beamformers, and infinitely many constraints caused by the imperfect CSI, the formulated problem is challenging to solve. For computational complexity reduction, the original problem is divided into a user grouping subproblem and a robust beamforming design subproblem. First, the user grouping problem is efficiently solved by a coalition formation game based algorithm. Then, for the robust beamforming design problem, a semidefinite relaxation (SDR) based method is proposed to produce a suboptimal solution efficiently. Moreover, we provide a sufficient condition under which the SDR based approach can guarantee to obtain an optimal rank-one solution, which is theoretically analyzed. Simulation results demonstrate the efficacy of the proposed algorithms.

Keywords: User grouping · Robust beamforming design · Coalition formation game · Semidefinite relaxation

1 Introduction

To support the exponentially growing data traffic and the number of devices in future wireless networks, non-orthogonal multiple access (NOMA) was proposed

to serve multiple users simultaneously on the same resource block [1–3]. The key idea of NOMA is to combine the superposition coding at the transmitter and successive interference cancellation (SIC) at the receiver, as such the spectral efficiency of wireless systems can be significantly improved [4, 5]. Due to its capability to provide superior spectral efficiency and massive connectivity, NOMA has been applied to many aspects of wireless communication systems, e.g., the Internet of Things network [6], intelligent reflective surface network [7], and mobile edge computing network [8]. User grouping, also known as user pairing, is one fundamental issue of NOMA. The impacts of user grouping on the system performance have been intensively investigated from both performance analysis and system design perspectives. For example, [9] studied the influences of user pairing on the outage probabilities of users in the NOMA systems. [10] proposed to use a branch-and-bound based algorithm to solve the joint user pairing and power allocation optimally with the worst-case computation complexity of NP-hard. To develop a computation-efficient algorithm for the user grouping problem, matching theory based heuristic algorithms were presented in [10–12].

To benefit from additional spatial degrees of freedom, applying the multi-antenna technique to NOMA is an important way to further improve the energy and spectrum efficiency of the systems. To fully exploit the advantages of the multi-antenna technique, beamforming design relying on the channel state information (CSI) between the BS and users has been widely studied in the literature. For example, For more practical applications, the beamforming designs for multi-cell NOMA systems have been investigated in [13–15]. However, the above works are all based on the assumption that the CSI is perfectly known by the system. However, all the aforementioned works assumed that the CSI of the system can be perfectly acquired. In practice, due to, e.g., imperfect channel estimation and finite feedback or channel aging [16], the system can never have perfect CSI. Therefore, it is interesting to investigate the robust design of communication systems under imperfect CSI assumptions. The robust beamforming design has been widely studied in the orthogonal frequency division multiple access systems, and recently, has been extended to the NOMA system [17]. However, to the best of our knowledge, none of the existing works considered the robust beamforming design in multi-cell NOMA systems with interference channels.

In this paper, we study the joint user grouping and robust beamforming design problem of a downlink multi-cell NOMA network with imperfect CSI assumptions. The formulated problem is a mixed-integer programming problem. Due to the integer variables relevant to user grouping, coupling effects of the downlink beamformers, and the infinitely many constraints brought by the imperfect CSI, the resultant problem is challenging to solve. For computational complexity reduction, the original problem is divided into the user grouping subproblem and robust beamforming design subproblem. For the user grouping subproblem, we present a coalition formation game based algorithm to efficiently determine the user grouping. Under imperfect CSI assumption, the worse-case robust beamforming design problem, with signal-to-interference-noise-ratio (SINR) and SIC constraints is also investigated. The coupling effects of the beamformers, both from inter- and intra-group, make the formulated problem

nonconvex and challenging to solve. Moreover, the imperfect CSI assumption makes the considered problem much more complicated, since each of the worst-case SINR or SIC constraints corresponds to an infinite number of nonconvex constraints. The contributions of this work are summarized as follows:

- We formulate the joint user grouping and robust beamforming problem in a multi-cell downlink NOMA system under imperfect CSI assumptions. For computation-efficiency consideration, the original problem is decoupled into two subproblems, i.e., the user grouping problem and the robust beamforming problem. To solve the user grouping problem efficiently, we propose to use a coalition formation game based user grouping algorithm.
- For the robust beamforming design problem. To simplify the corresponding problem, we first use the semidefinite relaxation (SDR) based method to transform the quadratic terms to beamformers into linear ones, and then, the S-lemma is invoked to deal with the infinitely many constraints caused by the imperfect CSI. By omitting the rank-one constraints, the reformulated problem refers to the semidefinite programming (SDP) problem, which is convex and can be efficiently solved by the existing optimization tools. Further, to gain more insights into the proposed SDR-based algorithm, a sufficient condition, under which the rank-one optimality of the obtained solution can be guaranteed, is provided and the rank-one optimality of the obtained solution is theoretically proved.

2 System Model and Problem Formulation

Consider a downlink multi-cell network which consists of M multi-antenna BSs and K single-antenna users with $K \geq 2M$. Let $\mathcal{M} \triangleq \{1, 2, \dots, M\}$ and $\mathcal{K} \triangleq \{1, 2, \dots, K\}$ denote the index sets of BSs and users, respectively. The randomly deployed users are grouped into M non-overlapping user groups, i.e., $\mathcal{C}_n \cap \mathcal{C}_{\hat{n}} = \emptyset, \forall n \neq \hat{n}$ where $\mathcal{C}_n \triangleq \{U_{n1}, U_{n2}, \dots, U_{nN_n}\}$ denotes the user set of the n -th group and $N_n = |\mathcal{C}_n|$ is the number of users in the n -th group, satisfying $\sum_{n=1}^M N_n = K$. Each group is served by a BS.

Denote by, $\mathbf{h}_{mnk}^H \in \mathbb{C}^{N_t}$, the channel from BS m to U_{nk} . As discussed previously that the BSs inevitably suffer from CSI errors in practical systems. Thus, imperfect CSI model is considered in this work. Let $\mathbf{h}_{mnk} \in \mathbb{C}^{N_t}$ denote the pre-assumed CSI at the BS m for U_{nk} . Then, the real CSI between BS m and U_{nk} is given by

$$\mathbf{h}_{mnk} = \tilde{\mathbf{h}}_{mnk} + \mathbf{e}_{mnk}, \forall m, n \in \mathcal{M}, k \in \mathcal{C}_n, \quad (1)$$

where \mathbf{e}_{mnk} is the bounded CSI error associated with \mathbf{h}_{mnk} . In particular, the bounded CSI error can be modelled by

$$\mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1, \forall m, n \in \mathcal{M}, k \in \mathcal{C}_n, \quad (2)$$

where $\mathbf{Q} \in \mathbb{H}^{N_t}$ determines the range and shape of the CSI error. For instance, $\mathbf{Q} = \frac{1}{\epsilon^2} \mathbf{I}_{N_t}$ characterizes the popular spherical error model $\|\mathbf{e}_{mnk}\|^2 \leq \epsilon^2$.

For spectral efficiency consideration, the NOMA technique is applied to each user group. Following the rationale of NOMA, the signals of users in group n are combined by using the superposition coding technique, then the users with stronger channel conditions will first remove the signals for the users with weaker channel conditions by invoking the SIC technique. Let $s_{nk} \in \mathbb{C}$ denote the signal for user k in group n with $\mathbb{E}\{|s_{nk}|^2\} = 1$. So, after superposition coding, the transmit signal of BS n is given by

$$\mathbf{s}_n = \sum_{k=1}^{N_n} \mathbf{w}_{nk} s_{nk}, \forall n \in \mathcal{M}, \quad (3)$$

where \mathbf{w}_{nk} is the beamformer for U_{nk} . The received signal at U_{nk} is give by

$$y_{nk} = \mathbf{h}_{nnk}^H \mathbf{w}_{nk} s_{nk} + \mathbf{h}_{nnk}^H \sum_{i \neq k}^{N_n} \mathbf{w}_{ni} s_{ni} + \sum_{m \neq n}^M \mathbf{h}_{mnk}^H \sum_{i=1}^{N_m} \mathbf{w}_{mi} s_{mi} + z_{nk}, \quad (4)$$

where $z_{nk} \in \mathbb{C}$ is the received additive white Gaussian noise at U_{nk} with zero mean and variance σ_{nk}^2 . The first item in (4) denotes the desired signal of U_{nk} , the second and third items in (4) denote the intra-cell and inter-cell interference, respectively.

Without loss of generality, we assume that the users, in each group, are ordered by their channel gains in a descending manner, i.e., $|\mathbf{h}_{mn1}|^2 \geq |\mathbf{h}_{mn2}|^2 \geq \dots \geq |\mathbf{h}_{mnN_n}|^2$. Thus, according to the principle of NOMA, U_{nk} would first remove the information, s_{nj} , for U_{nj} for $j > k$ by using SIC, and then decoding its own information. Based on the above model, the SINR for decoding s_{nj} , $\forall j > k$, at U_{nk} in the SIC process is given by

$$\begin{aligned} & \text{SINR}_{nk}^{s_{nj}} (\{\mathbf{w}_{nk}\}_{\forall n,k}, \{\mathbf{h}_{mnk}\}_{\forall m,n,k}) \\ &= \frac{|\mathbf{h}_{nnk}^H \mathbf{w}_{nj}|^2}{\sum_{i=1}^{j-1} |\mathbf{h}_{nnk}^H \mathbf{w}_{ni}|^2 + \sum_{m \neq n}^M \sum_{i=1}^{N_m} |\mathbf{h}_{mnk}^H \mathbf{w}_{mi}|^2 + \sigma_{nk}^2}, \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1. \end{aligned} \quad (5)$$

After s_{nj} is removed from the received signal, the SINR at U_{nk} for decoding s_{nk} is given by

$$\begin{aligned} & \text{SINR}_{nk}^{s_{nk}} (\{\mathbf{w}_{nk}\}_{\forall n,k}, \{\mathbf{h}_{mnk}\}_{\forall m,n,k}) \\ &= \frac{|\mathbf{h}_{nnk}^H \mathbf{w}_{nk}|^2}{\sum_{i=1}^{k-1} |\mathbf{h}_{nnk}^H \mathbf{w}_{ni}|^2 + \sum_{m \neq n}^M \sum_{i=1}^{N_m} |\mathbf{h}_{mnk}^H \mathbf{w}_{mi}|^2 + \sigma_{nk}^2}, \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1. \end{aligned} \quad (6)$$

Based on the above model, the energy-efficient joint user-grouping and robust beamforming design problem can be formulated as

$$\min_{\{C_n\}_{n=1}^M, \{\mathbf{w}_{nk}\}_{n,k}} \sum_{n=1}^M \sum_{k=1}^{N_n} \|\mathbf{w}_{nk}\|^2 \quad (7a)$$

$$\text{s.t. } \mathcal{C}_n \cap \mathcal{C}_{n'} = \emptyset, \forall n \neq n' \quad (7b)$$

$$\begin{aligned} \text{SINR}_{nk}^{s_{nj}}(\{\mathbf{w}_{nk}\}_{n,k}, \{\mathbf{h}_{mnk}\}_{m,n,k}) &\geq \gamma_{nj}, \\ \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} &\leq 1, m, n \in \mathcal{M}, k \in \mathcal{C}_n, j > k, \end{aligned} \quad (7c)$$

$$\begin{aligned} \text{SINR}_{nk}^{s_{nk}}(\{\mathbf{w}_{nk}\}_{n,k}, \{\mathbf{h}_{mnk}\}_{m,n,k}) &\geq \gamma_{nk}, \\ \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} &\leq 1, m, n \in \mathcal{M}, k \in \mathcal{C}_n, \end{aligned} \quad (7d)$$

$$\sum_{k=1}^{N_n} \|\mathbf{w}_{nk}\|^2 \leq P_{\max}, \forall n \in \mathcal{M}, \quad (7e)$$

where P_{\max} is the transmission power budget of BS, (7b) represents that each user will be uniquely assigned into one group, (7c) guarantees the success of SIC procedure at each user, and (7e) signifies the Quality-of-Service requirement for each user.

It is not difficult to verify that problem (7) is a nonconvex optimization problem due to the coupling of the quadratic beamforming vectors and also the channel uncertainty. To efficiently produce a high-quality solution to problem (7), similar to [11, 12, 18], we will decouple it into two subproblems, i.e., the user grouping problem and the robust beamforming design problem. In the following sections, the user paring problem and the robust beamforming design problem will be solved alternatively. Specifically, the user paring problem is solved by a coalition formation game based algorithm and the robust beamforming design problem is handled by the semidefinite relaxation based algorithm.

3 Coalition Formation Game Based Algorithm for User Grouping

In this section, we solve the user paring problem via the coalition formation game based algorithm. To this end, we first introduce the definition of a coalition formation game.

Definition 1 (Coalition Formation Game [19]). A coalition formation game is given by a pair $(\mathcal{K}, \mathcal{P})$, where \mathcal{K} is the set of players, and $\mathcal{P} = [\succ_1, \succ_2, \dots, \succ_K]$ denotes the preference profiles, specifying for each player $k \in \mathcal{K}$ its preference relation \succ_k .

In this work, the mobile users are viewed as players, and BSs are viewed as coalitions. The set of players is $\mathcal{K} = \{U_1, U_2, \dots, U_K\}$. For the sake of distinction, denote $\mathcal{C} = \{C_1, C_2, \dots, C_M\}$ as the set of coalitions. Users who are associated with the same BS are referred to as players who join the same coalition. We assume that players are more willing to join the coalition yielding higher social welfare. In other words, they are cooperative rather than rational (selfish).

At the beginning of coalition formation, each player proposes to joint a coalition according to its preference list which is built based on the consumed transmit power and can be described as follows:

$$U_{PF}(k) = [C_C(1), C_C(2), \dots, C_C(j), \dots, C_C(K)], \forall k \in \mathcal{K}, \quad (8)$$

where the lists are sorted in an ascending order of the consumed power. The preference lists satisfy the condition that the BS who can provide the minimum consumed power in the coalition set will be ranked as the first one. This can be described as

$$U(k)_j \succ_k U(k)_{j'}, \forall k \in \mathcal{K} \quad (9)$$

for all k, j in \mathcal{K} and \mathcal{C} respectively. This means that player k prefers to form coalition with BS j to j' if the total consumed power of user pair $\langle U_i, C_j \rangle$ is less than that of the user pair $\langle U_i, C_{j'} \rangle$.

To better handle the minimization of the total transmit power, we introduce the swap operation between two players, e.g., U_i and $U_{i'}$, in \mathcal{K} and the corresponding two matched coalitions, e.g., $\mathcal{F}(U_{i'})$ and $\mathcal{F}(U_i)$, in \mathcal{C} in the coalition formation process. The swap operation can be defined by

$$\begin{aligned} \mathcal{F}_i^{i'} &= \mathcal{F} \setminus \{ \langle U_i, \mathcal{F}(U_i) \rangle, \langle U_{i'}, \mathcal{F}(U_{i'}) \rangle \} \\ &\cup \{ \langle U_i, \mathcal{F}(U_{i'}) \rangle, \langle U_{i'}, \mathcal{F}(U_i) \rangle \}, \end{aligned} \quad (10)$$

where U_i and $U_{i'}$ switch the matched BSs while keeping other coalitions invariant. Based on the swap operation, we define the swap blocking pair as below:

Definition 2 (swap blocking pair). Given a matching \mathcal{F} and two user pairs $\langle U_i, C_j \rangle$ and $\langle U_{i'}, C_{j'} \rangle$. If there exists a new coalition formation $\mathcal{F}_i^{i'}$ such that the total transmit power of the new user pairs gets a decrease, then the swap operation is approved, and $\langle U_i, C_{j'} \rangle, \langle U_{i'}, C_j \rangle$ are swap blocking pairs under the coalition formation $\mathcal{F}_i^{i'}$.

Note that the above definition implies that there is a benefit by exchanging the coalition formation of $\langle U_i, C_j \rangle, \langle U_{i'}, C_{j'} \rangle$ and this operation will not hurt the benefit of other coalitions. Thus, with this new coalition formation, the total transmission power of the system can be decreased. Based on the above definitions, the coalition formation procedure can be described as follows. Firstly, each user can be associated with a BS to form a potential swap blocking pair. Then the system will check whether these two user pairs can get a benefit by exchanging their current coalition formation. The users will keep performing approved swap operations until they reach a stable status, which is known as *Nash stable*. Its definition is described below.

Definition 3 (Nash-stable [20]). A coalition partition is Nash-stable if and only if the following formula holds:

$$U(k)_i \succ_k U_{\mathcal{K}}(k)_j, \forall k \in \mathcal{K}, i, j \in \mathcal{C}, i \neq j, \quad (11)$$

which implies that, in a Nash-stable partition, any player is more (at least as well) willing to stay in its current coalition than to join other coalitions. In other words, there does not exist such a coalition that any player strictly prefers over its current one.

As per the above definitions, we can develop a coalition formation game based algorithm to solve the user pairing problem. The detailed procedure is summarized in Algorithm 1. It is not difficult to verify that the coalition formation game based algorithm can converge to a Nash-stable partition.

Algorithm 1. Coalition formation game based algorithm for user grouping

1: Initialize preference lists for players as $U_{PF}(i), i \in \mathcal{K}$.

Stage I : initial coalition formation

2: **for** $i = 1$ to K **do**

3: Each player U_i requests to its preferred BS j according to $U_{PF}(i)$.

4: **if** The number of joined players does not exceed the maximum of C_j **then**

5: Player U_i joins the coalition formation of C_j .

6: **else**

7: U_i will join the other coalition according to $U_{PF}(i)$.

8: **end if**

9: **end for**

Stage II : swap operation

10: **repeat**

11: Search the coalition formation to check whether there exists swap blocking pair.

12: **if** there is swap blocking pair **then**

13: Swap the user pair and update the current coalition formation.

14: **end if**

15: **until** There is no swap blocking pair in the coalition formation.

The computational complexity of Algorithm 1 consists of two parts. One is due to the initial coalition formation phase, which has a complexity order of $\mathcal{O}(K)$. The other is due to the swap operations in the second phase. For each user, there exist $M - 1$ possible BSs to do swapping, thus the complexity order is given by $\mathcal{O}(K(M - 1))$. Therefore, the total complexity is $\mathcal{O}(MK)$. Compared to the optimal strategy using exhaustive search, which has a complexity order of $\mathcal{O}(M^K)$, the computational complexity of the proposed coalition formation game based algorithm is dramatically decreased.

4 Solve the Robust Beamforming Design Problem

Once the user grouping is determined, problem (7) boils down to the robust beamforming design problem. In this section, we will propose to use the SDR method to solve the beamforming design problem, and the optimality of the obtained solution is analyzed.

4.1 SDR-Based Algorithm Design

To apply the SDR method, we first introduce a set of rank-one matrix $\mathbf{W}_{n,k} = \mathbf{w}_{nk} \mathbf{w}_{nk}^H, \forall n, k$. Then, by ignoring the rank-one constraint on the matrix, the robust beamforming problem can be relaxed as

$$\min_{\{\mathbf{W}_{nk}\}_{\forall n,k}} \sum_{n=1}^M \sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \quad (12a)$$

$$\text{s.t. } \mathbf{h}_{nnk}^H \left(\frac{1}{\gamma_{nj}} \mathbf{W}_{nj} - \sum_{i=1}^{j-1} \mathbf{W}_{nk} \right) \mathbf{h}_{nnk} \geq \sum_{m \neq n}^M \mathbf{h}_{mnk}^H \left(\sum_{i=1}^{N_m} \mathbf{W}_{mi} \right) \mathbf{h}_{mnk} + \sigma_{nk}^2, \\ \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1, m, n, j > k, \quad (12b)$$

$$\mathbf{h}_{nnk}^H \left(\frac{1}{\gamma_{nk}} \mathbf{W}_{nk} - \sum_{i=1}^{k-1} \mathbf{W}_{ni} \right) \mathbf{h}_{nnk} \geq \sum_{m \neq n}^M \mathbf{h}_{mnk}^H \left(\sum_{i=1}^{N_m} \mathbf{W}_{mi} \right) \mathbf{h}_{mnk} + \sigma_{nk}^2, \\ \forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1, m, n, k, \quad (12c)$$

$$\sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \leq P_{\max}, \forall n, \quad (12d)$$

$$\mathbf{W}_{nk} \succeq \mathbf{0}, \forall n, k, \quad (12e)$$

which is a convex problem as the objective function and constraints are linear. However, it is still computationally intractable due to the infinite number of constraints. Next, to make problem (12) tractable, we propose the following lemma.

Lemma 1. *The infinitely many constraints in (12b) and (12c) can be equivalently recast into the following finite number of linear matrix inequalities:*

$$\Phi_{nj} \left(\{\mathbf{W}_{ni}\}_{i=1}^j, \{\theta_{mnk}\}_m, \lambda_{mnk} \right) \\ \triangleq \begin{bmatrix} \mathbf{A}_{nj} + \lambda_{nnk} \mathbf{Q}_{nnk} & \mathbf{A}_{nj} \hat{\mathbf{h}}_{nnk} \\ \hat{\mathbf{h}}_{nnk}^H \mathbf{A}_{nj} & \hat{\mathbf{h}}_{nnk}^H \mathbf{A}_{nj} \hat{\mathbf{h}}_{nnk} - \sum_{m \neq n}^M \theta_{mnk} - \sigma_{nk}^2 - \lambda_{nnk} \end{bmatrix} \succeq \mathbf{0}, \forall m, n, j > k, \quad (13a)$$

$$\Psi_{nk} \left(\{\mathbf{W}_{ni}\}_{i=1}^k, \{\theta_{mnk}\}_m, \lambda_{mnk} \right) \\ \triangleq \begin{bmatrix} \mathbf{B}_{nk} + \lambda_{nnk} \mathbf{Q}_{nnk} & \mathbf{B}_{nk} \hat{\mathbf{h}}_{nnk} \\ \hat{\mathbf{h}}_{nnk}^H \mathbf{B}_{nk} & \hat{\mathbf{h}}_{nnk}^H \mathbf{B}_{nk} \hat{\mathbf{h}}_{nnk} - \sum_{m \neq n}^M \theta_{mnk} - \sigma_{nk}^2 - \lambda_{nnk} \end{bmatrix} \succeq \mathbf{0}, \forall m, n, k, \quad (13b)$$

$$\Omega_{mnk} \left(\{\mathbf{W}_{ni}\}_{i=1}^{N_n}, \theta_{mnk}, \lambda_{mnk} \right) \\ \triangleq \begin{bmatrix} -\mathbf{C}_m + \lambda_{mnk} \mathbf{Q}_{mnk} & -\mathbf{C}_m \hat{\mathbf{h}}_{mnk} \\ -\hat{\mathbf{h}}_{mnk}^H \mathbf{C}_m & -\hat{\mathbf{h}}_{mnk}^H \mathbf{C}_m \hat{\mathbf{h}}_{mnk} + \theta_{mnk} + \lambda_{mnk} \end{bmatrix} \succeq \mathbf{0}, \forall m, n, k, \quad (13c)$$

where $\lambda_{mnk} > 0, \forall m, n \in \mathcal{M}, k \in \mathcal{C}_n$ are auxiliary variables and

$$\begin{aligned} \mathbf{A}_{nj} &= \frac{1}{\gamma_{nj}} \mathbf{W}_{nj} - \sum_{i=1}^{j-1} \mathbf{W}_{ni}, & \mathbf{B}_{nk} &= \frac{1}{\gamma_{nk}} \mathbf{W}_{nk} - \sum_{i=1}^{k-1} \mathbf{W}_{ni}, \\ \mathbf{C}_m &= \sum_{i=1}^{N_m} \mathbf{W}_{mi}, \\ \theta_{mnk} &= \max_{\mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1} \mathbf{h}_{mnk}^H \left(\sum_{i=1}^{N_m} \mathbf{W}_{mi} \right) \mathbf{h}_{mnk}. \end{aligned}$$

Proof. The key idea to prove Lemma 1 is using S-lemma to handle the infinitely many constraints. We omit it here due to the space limitation.

Based on Lemma 1, the SDP problem (12) can be equivalently reformulated as

$$\min_{\{\mathbf{W}_{nk}\}, \{\theta_{mnk}\}, \{\lambda_{mnk}\}} \sum_{n=1}^M \sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \quad (14a)$$

$$\text{s.t. } \Phi_{nj} \left(\{\mathbf{W}_{ni}\}_{i=1}^j, \{\theta_{mnk}\}_m, \lambda_{nnk} \right) \succeq \mathbf{0}, \forall m, n, j > k, \quad (14b)$$

$$\Psi_{nk} \left(\{\mathbf{W}_{ni}\}_{i=1}^k, \{\theta_{mnk}\}_m, \lambda_{nnk} \right) \succeq \mathbf{0}, \forall m, n, k, \quad (14c)$$

$$\Omega_{mnk} \left(\{\mathbf{W}_{ni}\}_{i=1}^{N_n}, \theta_{mnk}, \lambda_{mnk} \right) \succeq \mathbf{0}, \forall m, n, k, \quad (14d)$$

$$\sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \leq P_{\max}, \forall n, \quad (14e)$$

$$\mathbf{W}_{nk} \succeq \mathbf{0}, \lambda_{mnk} \geq 0, \forall n, k, \quad (14f)$$

which is an SDP and thus can be efficiently solved by CVX.

Remind that problem (12) is a relaxed version of the original robust beamforming design problem by ignoring the rank-one constraints. Thus, one important issue in solving problem (12) is to verify whether the obtained matrices from solving problem (14) are rank-one. If it is true, then the optimal beamforming vectors can be obtained by simply applying singular value decomposition to the obtained matrices. Hence, it is interesting to explore the conditions under which solving problem (14) can produce rank-one solutions.

4.2 Rank-One Optimality Analysis

The following Lemma provides a condition that can guarantee the rank optimality of problem (14).

Lemma 2. *Suppose that problem (14) is feasible, the rank-one optimality can be guaranteed if $\mathbf{Q}_{nnk} = \infty \mathbf{I}_{N_c}$ for all n, k , i.e., no intra-cell CSI error.*

Proof. As perfect intra-cell CSI errors are available at the BSs, problem (12) degrades to

$$\min_{\{\mathbf{W}_{nk}\}_{\forall n,k}} \sum_{n=1}^M \sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \quad (15a)$$

$$\begin{aligned} \text{s.t. } & \hat{\mathbf{h}}_{nnk}^H \left(\frac{1}{\gamma_{nj}} \mathbf{W}_{nj} - \sum_{i=1}^{j-1} \mathbf{W}_{nk} \right) \hat{\mathbf{h}}_{nnk} \\ & \geq \max_{\forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1} \left\{ \sum_{m \neq n}^M \hat{\mathbf{h}}_{mnk}^H \left(\sum_{i=1}^{N_m} \mathbf{W}_{mi} \right) \hat{\mathbf{h}}_{mnk} \right\} + \sigma_{nk}^2, \forall n, k, \end{aligned} \quad (15b)$$

$$\begin{aligned} & \hat{\mathbf{h}}_{nnk}^H \left(\frac{1}{\gamma_{nk}} \mathbf{W}_{nk} - \sum_{i=1}^{k-1} \mathbf{W}_{ni} \right) \hat{\mathbf{h}}_{nnk} \\ & \geq \max_{\forall \mathbf{e}_{mnk}^H \mathbf{Q}_{mnk} \mathbf{e}_{mnk} \leq 1} \left\{ \sum_{m \neq n}^M \hat{\mathbf{h}}_{mnk}^H \left(\sum_{i=1}^{N_m} \mathbf{W}_{mi} \right) \hat{\mathbf{h}}_{mnk} \right\} + \sigma_{nk}^2, \forall n, k, \end{aligned} \quad (15c)$$

$$\sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \leq P_{\max}, \forall n, k, \quad (15d)$$

$$\mathbf{W}_{nk} \succeq \mathbf{0}, \forall n, k. \quad (15e)$$

Again, by applying *S-Lemma* to the right-hand-sides of (15b) and (15c), problem (15) can be equivalently reformulated as:

$$\min_{\{\mathbf{W}_{nk}\}, v_{mnk}, \lambda_{mnk}} \sum_{n=1}^M \sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \quad (16a)$$

$$\text{s.t. } \hat{\mathbf{h}}_{nnk}^H \left(\frac{1}{\gamma_{nj}} \mathbf{W}_{nj} - \sum_{i=1}^{j-1} \mathbf{W}_{nk} \right) \hat{\mathbf{h}}_{nnk} \geq \sum_{m=1, m \neq n} v_{mnk} + \sigma_{nk}^2, \forall n, k, \quad (16b)$$

$$\hat{\mathbf{h}}_{nnk}^H \left(\frac{1}{\gamma_{nk}} \mathbf{W}_{nk} - \sum_{i=1}^{k-1} \mathbf{W}_{ni} \right) \hat{\mathbf{h}}_{nnk} \geq \sum_{m=1, m \neq n} v_{mnk} + \sigma_{nk}^2, \forall n, k, \quad (16c)$$

$$\begin{bmatrix} -\sum_{i=1}^{N_m} \mathbf{W}_{mi} + \lambda_{mnk} \mathbf{Q}_{mnk} & -\sum_{i=1}^{N_m} \mathbf{W}_{mi} \hat{\mathbf{h}}_{mnk} \\ -\hat{\mathbf{h}}_{mnk}^H \sum_{i=1}^{N_m} \mathbf{W}_{mi} - \hat{\mathbf{h}}_{mnk}^H \sum_{i=1}^{N_m} \mathbf{W}_{mi} \hat{\mathbf{h}}_{mnk} + v_{mnk} + \lambda_{mnk} \end{bmatrix} \succeq \mathbf{0}, \forall n, k, \quad (16d)$$

$$\sum_{k=1}^{N_n} \text{Tr}(\mathbf{W}_{nk}) \leq P_{\max}, \forall n, k, \quad (16e)$$

$$\mathbf{W}_{nk} \succeq \mathbf{0}, \forall n, k. \quad (16f)$$

We prove Lemma 2 by using the KKT conditions of (16). In particular, let $\{\delta_{nj}^*\}$, $\{\epsilon_{nj}^*\}$, \mathbf{Y}_{mi}^* and \mathbf{Z}_{nk}^* denote the optimal dual variables associated with

(16b), (16c), (16d) and (16f), respectively. The KKT conditions related to \mathbf{W}_{nk}^* are as follows:

$$\mathbf{Z}_{nk}^* \mathbf{W}_{nk}^* = \mathbf{0}, \quad (17a)$$

$$\mathbf{Z}_{nk}^* = \mathbf{I}_{N_t} - \left(\frac{\delta_{nk}^*}{\gamma_{nj}} + \frac{\epsilon_{nk}^*}{\gamma_{nk}} \right) \hat{\mathbf{h}}_{nnk} \hat{\mathbf{h}}_{nnk}^H, \quad (17b)$$

$$\mathbf{Z}_{nk}^* \succeq \mathbf{0}, \mathbf{W}_{nk}^* \succeq \mathbf{0} \quad (17c)$$

First, note that, from (16c), we can conclude that $\mathbf{W}_{nk}^* \succ \mathbf{0}$. Otherwise, we have

$$-\hat{\mathbf{h}}_{nnk}^H \sum_{i=1}^{k-1} \mathbf{W}_{ni} \hat{\mathbf{h}}_{nnk} \geq \sum_{m=1, m \neq n} v_{mnk} + \sigma_{nk}^2,$$

which violates the fact that $v_{mnk} \geq 0$, $\sigma_{nk}^2 > 0$ and \mathbf{W}_{ni} are positive semidefinite. Then, also note that

$$0 = \text{rank}(\mathbf{Z}_{nk}^* \mathbf{W}_{nk}^*) \geq \text{rank}(\mathbf{Z}_{nk}^*) + \text{rank}(\mathbf{W}_{nk}^*) - N_t. \quad (18)$$

So, we have $\text{rank}(\mathbf{W}_{nk}^*) \leq N_t - \text{rank}(\mathbf{Z}_{nk}^*)$. Therefore, to prove that \mathbf{W}_{nk}^* is rank-one, it suffices to prove that

$$\begin{aligned} \text{rank}(\mathbf{Z}_{nk}^*) &= \text{rank} \left(\mathbf{I}_{N_t} - \left(\frac{\delta_{nk}^*}{\gamma_{nj}} + \frac{\epsilon_{nk}^*}{\gamma_{nk}} \right) \hat{\mathbf{h}}_{nnk} \hat{\mathbf{h}}_{nnk}^H \right) \\ &= N_t - 1. \end{aligned} \quad (19)$$

Notice that $\hat{\mathbf{h}}_{nnk}$ is a non-zero vector, thus $\text{rank}(\hat{\mathbf{h}}_{nnk} \hat{\mathbf{h}}_{nnk}^H) = 1$. So, we have $\text{rank}(\mathbf{Z}_{nk}^*) \geq N_t - 1$. Finally, based on (18) and $\mathbf{W}_{nk}^* \succ \mathbf{0}$, we can conclude that $\text{rank}(\mathbf{Z}_{nk}^*) = N_t - 1$ and $\text{rank}(\mathbf{W}_{nk}^*) = 1$. This completes the proof.

We note that, for a general setup, the solutions of problem (14) *may not* be rank-one. In these cases, one can resort to the Gaussian randomization method to produce approximated beamforming vectors based on the obtained non-rank-one solutions [21].

5 Simulation Results

In this section, numerical simulations are present to verify the performance of the proposed transmission schemes and algorithms. The large-scale path loss is set to be $PL_{mnk} = 128.1 + 37.6 \log_{10}(d_{mnk})$ with d_{mnk} (in Km) denoting the distance between the BS m and U_{nk} , while the small-scale fading is modelled by the standard Rayleigh fading. The White noise power density is -174 dBm and the bandwidth is 10 MHz. The power budget of each BS is set as $P_{\max} = 36$ dBm. The users are randomly deployed in a $500 \text{ m} \times 500 \text{ m}$ square area. Without loss of generality, the QoS requirements of the users are the same, i.e., $\gamma_{mnk} = \gamma, \forall m, n \in \mathcal{M}, k \in \mathcal{C}_n$. The spherical error model

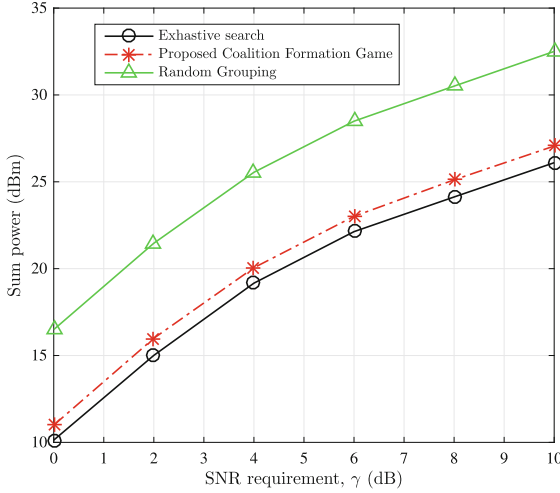


Fig. 1. The performance comparison of the centralized and decentralized algorithm with $M = 4$, $N = 3$, $N_t = 4$, $\epsilon = 0.05$ and $\gamma = 3$ dB.

is used and we assume that the error bounds of all users are the same, i.e., $\mathbf{Q}_{mnk} = \mathbf{Q} = \frac{1}{\epsilon^2} \mathbf{I}_{N_t}, \forall m, n \in \mathcal{M}, \forall k \in \mathcal{C}_n$. For comparison, we introduce some other user grouping approaches and transmission schemes, namely, the exhaustive search based user grouping approach, OMA transmission scheme, and non-robust NOMA transmission scheme, which are described as follows:

- **Exhaustive Search Based User grouping Approach:** In this approach, all the possible combinations of the user group are considered, the optimal user grouping is the one yielding the smallest transmit power.
- **Non-robust NOMA transmission scheme:** In this scheme, the perfect CSI of each user is assumed to be available at the BSs. It is not difficult to verify that the non-robust beamforming problem can be formulated as a SOCP and thus can be solved optimally.

We first evaluate the performance of the proposed coalition formation game based user grouping approach in Fig. 1. The exhaustive search based optimal user grouping algorithm is used as the performance benchmark. As it can be seen that the proposed coalition formation game based approach can achieve near-optimal performance under the system settings. While the random grouping based approach, in which the users are randomly selected to form a group, yields the worst performance.

In Fig. 2, we compare the performance of robust and non-robust designs versus different QoS requirements of users. It can be observed that as a price for a worst-case performance guarantee, the robust designs require higher transmission power than the non-robust design. From Fig. 3, we can also find that, with different CSI error bounds, the non-robust design would underestimate the required power for reliable transmissions. Meanwhile, the sum power increases

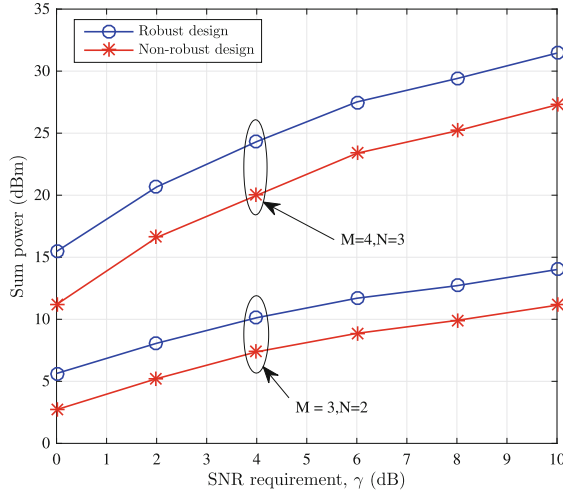


Fig. 2. The performance comparison of robust and non-robust designs versus different γ with $N_t = 4$, and $\epsilon = 0.05$.

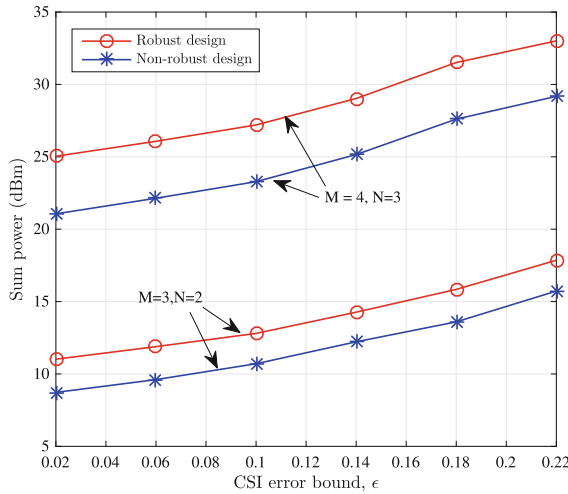


Fig. 3. The performance comparison of robust and non-robust designs versus different ϵ with $N_t = 4$, and $\gamma = 3$ dB.

with the increase of the CSI error bound. The reason is that a larger CSI error requires more transmission power to guarantee the QoS requirements of users.

6 Conclusions

In this work, we have investigated the joint user grouping and robust beamforming design problem in a downlink multi-cell NOMA network under imperfect

CSI assumptions. The formulated sum power minimization problem was shown to be challenging to solve due to the integer variables of user grouping, coupling effects of beamformers, and infinitely many constraints brought by the imperfect CSI. For computational complexity reduction, the original problem was decoupled into a user grouping subproblem and robust beamforming subproblem. A coalition formation game based algorithm has been proposed to solve the user grouping problem. While for the robust beamforming design problem, we proposed an SDR-based suboptimal algorithm was presented and a sufficient condition under which the SDR-based approach could guarantee to produce an optimal solution was presented. Simulation results have provided some interesting results. For example, the proposed coalition game based algorithm can perform close to the exhaustive search based scheme; The robust design would require more transmission power compared to the non-robust design.

Acknowledgements. The work of Y. Xu was supported by the China Postdoctoral Science Foundation under Grants No. 2021M693100 and Guangdong Basic and Applied Basic Research Foundation under Grants No. 2021A1515110018. The work of D. Cai was supported by National Natural Science Foundation of China (NSFC) under Grant No. 62001190 and the China Postdoctoral Science Foundation under Grants No. 2021M691249. The work of Z. Dong was supported by National Science Foundation of China under Grant 61561046 and the Key Research & Development and Transformation Plan of Science and Technology Program for Tibet Autonomous Region (No. XZ201901-GB-16).

References

1. Forecast, C.V.: Cisco visual networking index: global mobile data traffic forecast update, 2016–2017. White Paper, Cisco San Jose, CA, USA (2017)
2. Elbayoumi, M., Kamel, M., Hamouda, W., Youssef, A.: NOMA-assisted machine-type communications in UDN: state-of-the-art and challenges. *IEEE Commun. Surv. Tutor.* **22**, 1 (2020)
3. Ding, Z., Lei, X., Karagiannidis, G.K., Schober, R., Yuan, J., Bhargava, V.K.: A survey on non-orthogonal multiple access for 5G networks: research challenges and future trends. *IEEE J. Sel. Areas Commun.* **35**(10), 2181–2195 (2017)
4. Fang, F., Zhang, H., Cheng, J., Leung, V.C.M.: Energy-efficient resource allocation for downlink non-orthogonal multiple access network. *IEEE Trans. Commun.* **64**(9), 3722–3732 (2016)
5. Xu, Y., et al.: Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems. *IEEE Trans. Signal Process.* **15**(18), 4874–4886 (2017)
6. Wang, J., Kang, X., Sun, S., Liang, Y.: Throughput maximization for peer-assisted wireless powered IoT NOMA networks. *IEEE Trans. Wireless Commun.* **19**, 1 (2020)
7. Fang, F., Xu, Y., Pham, Q.-V., Ding, Z.: Energy-efficient design of IRS-NOMA networks. *IEEE Trans. Veh. Techn.* **69**(11), 14 088–14 092 (2020)
8. Fang, F., Xu, Y., Ding, Z., Shen, C., Peng, M., Karagiannidis, G.K.: Optimal resource allocation for delay minimization in NOMA-MEC networks. *IEEE Trans. Commun.* **68**(12), 7867–7881 (2020)

9. Ding, Z., Fan, P., Poor, H.V.: Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions. *IEEE Trans. Veh. Techn.* **65**(8), 6010–6023 (2016)
10. Cui, J., Liu, Y., Ding, Z., Fan, P., Nallanathan, A.: Optimal user scheduling and power allocation for millimeter wave NOMA systems. *IEEE Trans. Wireless Commun.* **17**(3), 1502–1517 (2018)
11. Di, B., Song, L., Li, Y.: Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks. *IEEE Trans. Wireless Commun.* **15**(11), 7686–7698 (2016)
12. Xu, Y., Cai, D., Fang, F., Ding, Z., Shen, C., Zhu, G.: Outage constrained power efficient design for downlink NOMA systems with partial HARQ. *IEEE Trans. Commun.* **68**(8), 5188–5201 (2020)
13. Ni, W., Liu, X., Liu, Y., Tian, H., Chen, Y.: Resource allocation for Multi-Cell IRS-Aided NOMA networks. *IEEE Trans. Wireless Commun.* **20**(7), 4253–4268 (2021)
14. Chen, G., Qiu, L., Ren, C.: On the performance of cluster-based MIMO-NOMA in Multi-Cell dense networks. *IEEE Trans. Commun.* **68**(8), 4773–4787 (2020)
15. Ding, J., Cai, J.: Two-side coalitional matching approach for joint MIMO-NOMA clustering and BS selection in Multi-Cell MIMO-NOMA systems. *IEEE Trans. Wireless Commun.* **19**(3), 2006–2021 (2020)
16. Cai, D., Xu, Y., Fang, F., Ding, Z., Fan, P.: On the impact of time-correlated fading for downlink NOMA. *IEEE Trans. Commun.* **67**(6), 4491–4504 (2019)
17. Wang, Z., Ng, D.W.K., Wong, V.W.S., Schober, R.: Robust beamforming design in C-RAN with sigmoidal utility and capacity-limited backhaul. *IEEE Trans. Wireless Commun.* **16**(9), 5583–5598 (2017)
18. Liang, W., Ding, Z., Li, Y., Song, L.: User pairing for downlink non-orthogonal multiple access networks using matching algorithm. *IEEE Trans. Commun.* **65**(12), 5319–5332 (2017)
19. Saad, W., Han, Z., Debbah, M., Hjørungnes, A., Basar, T.: Coalitional game theory for communication networks. *IEEE Signal Process. Mag.* **26**(5), 77–97 (2009)
20. Bogomolnaia, A., Jackson, M.O.: The stability of hedonic coalition structures. *Games Econ. Behav.* **38**(2), 201–230 (2002)
21. Luo, Z.Q., Ma, W.K., So, A.M.C., Ye, Y., Zhang, S.: Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.* **27**(3), 20–34 (2010)



The Vaccination Game in Susceptible-Infected-Susceptible (SIS) Networks with Multipopulations

Eitan Altman^{1,2,3(✉)}

¹ INRIA Sophia Antipolis - Méditerranée, Valbonne, France

eitan.altman@inria.fr

² LINCS, 75013 Paris, France

³ CERI/LIA, University of Avignon, Avignon, France

Abstract. We model in this paper the multipopulation vaccination game over a fully connected graph. Each player decides whether to purchase a vaccine or not, and if they do, then they further decide which vaccine to purchase among a finite number of vaccine producers. The players need not be indistinguishable. A potential consumer belongs to a risk type that characterizes how important it is for them to be vaccinated. The cost of a vaccine may depend on the demand, on the cost of the production, and on the consumer's class. We prove in the existence of an equilibrium within pure policies in the general multipopulation case. We further derive some properties of the equilibria in the case of a single risk-class.

Keywords: Vaccination game · Multipopulation

1 Introduction

Vaccination has turned out to be a major prevention method to fight both biological epidemics [14] as well as malware attacks [4, 8, 13]. Like many other viral pandemics, there is discrimination between the victims of the epidemics, and different infected persons may develop the disease in different forms or of degree of severity. But unlike other epidemics, getting infected with covid is correlated with belonging to some well defined risk groups such as sex, age group, other diseases etc. We call these groups sub-populations or risk classes. Members of each sub-population have similar properties and are assumed to be interchangeable. The knowledge of the above correlation has permitted decision makers to better treat higher risk groups and to avoid giving inappropriate treatments to specific risk groups.

In this vaccination game, each node in the network is a player faced with the decision of whether to buy a vaccine or not. We assume that any consumer may purchase any vaccine within some finite set of available candidates. If they buy a vaccine then they are immuned and do not infect themselves, nor other players.

We show that this game is equivalent to a crowding game (but not to a congestion game, as is the case for a single population [15]), and has therefore an equilibrium within pure strategies. We then identify the structure of equilibrium policies.

The structure of the paper is as follows. After a short Introduction, we present the model in Sect. 2. We recall the mean field approximation of the health quasi stationary probability distribution and model the health and the risk states. We then model the actions and the costs as a function of each vaccine and of the different risk-state (after averaging over the the health states that are assumed to be in the quasi stationary regime). The main results appear in Sect. 3. Further structural characterisation are derived for the special case of one population in Sect. 4. A concluding Section ends this paper.

2 The Model

2.1 Background: Quasi-stationary Infection Probability

We consider the Susceptible-Infected-Susceptible (SIS) compartment model on a non-directed graph. A node on this graph represents a (potential) consumer of a vaccine. The total number of nodes is denoted by N . The total number of consumers that decide to purchase the vaccine is given by n_v .

Let $V_i(t)$ denote the event that node i is infected at time t . S and I stand for the health state of a node: I is infected and S stands for Susceptible. $v_i(t)$ is the expectation of the indicator of (or simply the probability of) this event and A_{ij} is the incidence matrix. Its value is 1 if there is a link between node j and node i . Otherwise its value is zero. We assume that $A_{ij} = A_{ji}$. In the malware application, it means that the links are bi-directional. In the applications to pandemics propagation, this means that a node i is a potential contact of node j if and only if node j is a potential contact of node i .

The following holds:

$$\frac{dv_i(t)}{dt} = -\delta v_i(t) + \beta E \left[(1 - V_i(t)) \sum_{j=1}^N A_{ij} V_j(t) \right] \tag{1}$$

where δ is the rate of healing at an infected node, and β is the rate at which a susceptible node becomes infected by a contact with an infected neighbour. For any nodes i and j , we have $E[V_i(t)V_j(t)] \geq v_i v_j$ [5]. Hence for all i ,

$$\frac{dv_i(t)}{dt} \leq -\delta v_i(t) + \beta(1 - v_i(t)) \sum_{j=1}^N A_{ij} v_j(t) \tag{2}$$

The NIMFA mean field approximation v' is defined as the solution of (2) when replacing the inequality with an equality. The vector v' is a bound on the quasi stationary vector v and it becomes tight as the initial number of nodes N tends to infinity, see [4, 9, 15].

Let $\tau = \beta/\delta$ be the virality constant.

Proposition 1. $v(t) = 0$ is a stationary solution of (2). Assume throughout that all nodes are interconnected; i.e. all entries of the incident matrix are 1. $v = 0$ is the only solution of (2) if $\tau < 1/N$. A second symmetric vector solution whose entries are denoted by v_∞ is obtained when $\tau > 1/N$. It is given by

$$v_\infty = \frac{\tau N - 1}{\tau N} = 1 - \frac{1}{\tau N} \tag{3}$$

and it is called the quasi stationary solution.

Proof. We obtain the quasi stationary solutions by equating the time derivative to 0. Indeed, we get

$$v_\infty = 1 - \frac{1}{1 + \tau A v_\infty}$$

which implies

$$v_\infty = 1 - \frac{1}{1 + \tau N v_\infty} = \frac{\tau N v_\infty}{1 + \tau N v_\infty}$$

We conclude that

$$1 = \frac{\tau N}{1 + \tau N v_\infty}$$

thus getting condition (3).

2.2 States and Actions

We assume that the system is already in its quasi stationary regime which implies that the health states S and I of an individual are not used as part of a feedback control.

There are several ways in which one could obtain a multipopulation model. For example, it could be through the topology of the underlying network or through different time scales: a set of states such that every two nodes in the set are quickly reachable from a node are said to be from the same sub-population [10,15]. Below we introduce another classification into risk types according to the cost. This is motivated by risk states in covid and we say that all nodes with a similar cost have the same risk state. Risk state may stand for genre, age, other health problems etc. = We consider competition both between producers of vaccines as well as between consumers of vaccines. There are M vaccine producers and K risk classes of consumers. In this paper we study the competition between consumers of various risk-types; the competition is induced by given cost functions of the vaccine producers. We do not study here how these functions are formed, which would lead to price formation issues.

The action in this game for a consumer is a number m between 1 to $M + 1$; action m then stands for purchasing a vaccine from producer m . In addition, action $M + 1$ stands for not purchasing any vaccine.

The state in this game is given by the combination of a local component k which stands for the risk state of the player, and a global component \mathbf{n} . \mathbf{n} is a matrix of dimension $M + 1 \times K$. $n(m, k)$ is the number of players of type k that choose vaccine of producer m , $m = 1, \dots, M$. Let $n(M + 1)$ be the number of consumers who do not vaccinate. Thus $n(M + 1) + n_v = n$.

2.3 The Cost

The cost of type m vaccine is given as a function $C(m, k, n(m))$. The cost is to be understood in a wide sense. It is not just the monetary cost but may include also the risk of complications; which is why the cost depends on the consumer type k . The cost of type m vaccine is assumed to be a function of the total number $n(m)$ of individuals that purchase type m vaccine.

Remark 1. [15] study a similar game but the model there restricts to a single type of consumer and of vaccine. the vaccination comes at a fixed cost C that does not depend on the demand for the vaccine. Note that (3) implies that v is monotone increasing.

A node that invests in a vaccine is assumed not only to acquire resistance against the virus but also not to carry the virus anymore and thus to stop being contagious. We assume that all n nodes are initially connected (the incident matrix has one in all links) and thus if the number of nodes who purchase a vaccine is n_v then we are left with a fully connected graph with $n(M+1)$ nodes.

3 Main Result: A Multipopulation Equilibrium in Pure Strategies

The multipopulation game with multi-vaccine types that we have defined can be seen to be equivalent to the congestion game defined in Milchteich [6]. This class of games is known today as crowding games where as the term "congestion game" is used for games as defined in [11]. A crowding game is concerned with a network of parallel links all with a common source and a common destination. Players are atomic and each has to decide what link in the parallel link network to choose. A player belongs to one of K types. The cost for routing an agent through a given link is allowed to depend on the total load on that link, but this cost function may vary depending on the player (or on the class it belongs to). On the other hand, the contribution of a player to congestion in a link depends only on the total load on the link, and not on the class it belongs to. Nor is it a function of the class k of any other player in the network. The vaccine game is indeed equivalent to a crowding game in which link m corresponds to buying a vaccine of type m and where the choice $m = M+1$ corresponds to not purchasing any vaccine.

From the equivalence of this game to the vaccine game, we conclude the main result:

Theorem 1. *The multipopulation game has an equilibrium within pure policies in the multipopulation game.*

4 Further Structure in the Single Population Case

4.1 Existence of a Potential

We next derive the structure of policies in the case of a single risk class. We delete the risk-class from the notation. The cost functions for any player to

purchase a type m vaccine is given by $C(m, n(m))$, it thus depends on the total number of consumers that choose vaccine m . This is seen to be equivalent to a congestion game which describes routing games between atomic players. There is a restriction in these games that the cost of choosing to route through a given path should only depend on how many agents use that route and should not depend on other characteristics which are agent specific. Under this condition as well as the monotonicity of C in its second argument, the game is known to have a potential [11]. The potential of this game is given by

$$pot(\mathbf{n}) = \sum_{m=1}^{M+1} pot(\mathbf{n}, m) \quad \text{where} \tag{4}$$

$$pot(\mathbf{n}, m) = \sum_{\ell=1}^{n(m)} C(m, n(m))$$

This allows us to conclude that any local minimum of the potential is an equilibrium in the original game [12]. This further implies that the game has an equilibrium in pure policies. Furthermore, the equilibrium can be obtained by solving best response algorithms in a finite number of iterations.

4.2 Threshold Equilibrium Policies

A map u from the set of states to the set of actions is said to be a multi-policy. An equilibrium is a multi-policy from which no unilateral deviation is profitable for any player.

We present conditions in this section for equilibria policies to be of threshold type.

Definition 1. *A threshold policy for a player with a threshold value (n, q) , where n is an integer number and q is a nonnegative real number smaller than 1, is a policy that vaccinates if the number of players who are vaccinated is lower than n , and it does not vaccinate if the number of vaccinated is above n . At n the individual randomizes so that with probability q it vaccinates.*

Definition 2. *A symmetric multistrategy is said to be of a threshold type with parameter (n, q) if each player uses a threshold policy with that parameter.*

We restrict here to single type m of vaccine and a single type k of consumers. Each player has to decide whether to buy a vaccine or not. We shall assume that $C(n) - v(n)$ is increasing in n . This is the case in particular when C is taken to be a constant (corresponding to the cost of a vaccination).

Theorem 2. *Assume that n is the smallest integer such that $v(n) \geq C(n)$, i.e. when n purchase one vaccination each, then the cost $C(n)$ per vaccination is lower than the cost for getting infected by the virus. Then $v(\ell) > C(\ell)$ for all $\ell > n$ and $v(\ell) < C(\ell)$ for all $\ell < n$. Hence there is a symmetric threshold equilibrium multipolicy with parameter (n, q) for some q .*

Sketch of proof: Denote by $z = n + q$ the threshold policy with parameter (n, q) . We observe that if all players use a symmetric threshold policy z then the best response policy $z' = n' + q'$ for a player, say player i , is a threshold policy, monotonely decreasing in z . The Proof then follows Tarski fixed point Theorem [16].

Remark 2. i) Note that the equilibrium need not be unique. Indeed, this is the case if the symmetric equilibrium is not in pure policies. ii) A detailed proof of Theorem 2 can be adapted from the proof of Theorem 1 in [2].

5 Conclusion

Many free riding phenomena and competition may occur during pandemics. For fixed policy of health authority (in case of biological viruses) or of regulation bodies (in the case of e-viruses), individuals may respond in a cooperative or a non cooperative way. At a higher level, there may or there may not be cooperation and coordination between regulation bodies or public health institutions. we have started working on modeling these phenomena so as to better provide incentives to pursue cooperative behavior. This is important in view of the fact that the Covid 19 pandemic has had very different impacts in different countries- (less than 5 victims per 100000 in some countries (China, New Zealand, Australia) and more than a hundred in others (USA, France, Italy). In this paper we have studied the source of non-cooperation in the use of vaccines.

We focused on the game between potential consumers of vaccines who may be tempted not to vaccinate themselves thus adopting a free riding behavior. We modeled further the impact of pricing strategies of the producers of the vaccines on the equilibrium. We established existence of an equilibrium within pure policies. We obtain further structure of equilibria for the case of a single risk class.

References

1. Altman, E., Avritzer, A., El-Azouzi, R., Menasche, D.S., de Aguiar, L.P.: Rejuvenation and the Spread of Epidemics in General Topologies. In: 6th International Workshop on Software Aging and Rejuvenation (WoSAR 2014), Naples, Italy. <hal-01087263> (2014)
2. Altman, E., Shimkin, N.: Individually optimal dynamic routing in a processor sharing system: stochastic game analysis, EE Pub No. 849 (1992). A later version can be found in Operations Research, pp. 776–784, 1998
3. Khouzani, M.H.R., Sarkar, S., Altman, E.: Saddle-point strategies in malware attack. IEEE J. Sel. Areas Commun. **30**(1), 31–43 (2012)
4. Van Mieghem, P.: The N-intertwined SIS epidemic network model. Computing **93**(2–4), 147–169 (2011)
5. Van Mieghem, P., et al.: Virus spread networking. IEEE/ACM Trans. Networking **17**(1), 1–14 (2009)
6. Milchtaich, I.: Congestion games with player-specic payoff functions. Games Econom. Behav. **13**, 111–124 (1996)

7. Ozkaya, M., Izgi, B.: Effects of the quarantine on the individuals' risk of Covid-19 infection: game theoretical approach. *Alexandria Eng. J.* **60**(4), 4157–4165 (2021). ISSN 1110–0168, <https://doi.org/10.1016/j.aej.2021.02.021>, (<https://www.sciencedirect.com/science/article/pii/S1110016821000922>)
8. Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A.: Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015)
9. Prasse, B., Van Mieghem, P.: Time-dependent solution of the NIMFA equations around the epidemic threshold. *J. Math. Biol.* **81**, 1299–1355 (2020). <https://doi.org/10.1007/s00285-020-01542-6>
10. Sundaresan, R., Altman, E., Akhil, P.T.: Epidemics in a multicomunity network. [arXiv:1902.05713](https://arxiv.org/abs/1902.05713)
11. Rosenthal, R.W.: A class of games possessing pure-strategy Nash equilibria. *Internat. J. Game Theory* **2**, 65–67 (1973)
12. Monderer, D., Shapley, L.S.: Potential Games. *Games Econ. Behav.* **14**(1), 124–143 (1996). ISSN 0899–8256
13. Rufino, V.Q., et al.: Beyond herd immunity against strategic attackers. *IEEE Access* **8**, 66365–66399 (2020). <https://doi.org/10.1109/ACCESS.2020.2983652>
14. Scott, B.: The smallpox eradication game. *Public Choice*, 130(1/2), 179–207 (2007). <https://www.jstor.org/stable/27698049>
15. Trajanovski, S., et al.: Decentralized protection strategies against SIS epidemics in networks. *IEEE Trans. Control Network Syst.* **2**(4), 406–419 (2015). <https://doi.org/10.1109/TCNS.2015.2426755>
16. Tarski, A.: A lattice-theoretic fixed point theorem and its applications. *Pac. J. Math.* **5**, 285–309 (1955)
17. Wang, J.J.: The president's dilemma: open-up versus lock-down amid COVID-19 - using game theory to study the optimal strategies to combat the pandemic. *J. Game Theory* **9**(1), 8–12 (2020). ISSN: 2325–0046 e-ISSN: 2325–0054. <https://doi.org/10.5923/j.jgt.20200901.02>



A Stochastic Bandwidth Scanning Game

Andrey Garnae^(✉) and Wade Trappe

WINLAB, Rutgers University, North Brunswick, USA
garnae^v@yahoo.com, trappe@winlab.rutgers.edu

Abstract. In this paper we consider a dilemma that arises in bandwidth scanning problems associated with the design of agents' scanning strategies based on the principle of rationality and the principle of insufficient reasons. On one hand, engaging tools that estimate a network's parameters allows an agent to act rationally to maximize its payoff. On the other hand, utilizing such engagement incurs extra costs associated with scanning. In particular, if the agent does not employ such tools, then the involved expenses can be reduced, although such a strategy might also cause a reduction in detection probability since in such cases the agent has to design strategy based on the principle of insufficient reasons (also called principle of indifference). In this paper we model this dilemma as a non-zero sum stochastic game between two players (Scanner and Invader). The equilibrium is found in closed form in stationary strategies via solving the corresponding Shapley-Bellman equations, and its dependence on network parameters is illustrated.

Keywords: Non-zero sum stochastic game · Detection · Equilibrium

1 Introduction

The open nature of the wireless medium, in spite of a lot of benefits given by the ability to access spectrum dynamically, also makes cognitive radios a powerful tool for conducting malicious activities or policy violations by secondary users. Therefore, detecting malicious users or unlicensed activities is a crucial problem facing dynamic spectrum access [15], and one of the challenges to enforcing the proper usage of spectrum is the development of an intrusion detection systems that can scan large amounts of spectrum and identify illegal activity [5]. Since, in such security problems, there are at least two agents with different goals (the adversary aims to sneak into bands undetected for their illegal usage, while the intrusion detection system intends to prevent illegal spectrum usage), *game theory* is an ideal tool to employ. As examples of applying game theory to detect an adversary to prevent malicious attack on networks, we mention [1–3, 6, 10–12, 16, 19–22, 24]. In all of these papers the players were *rational*, i.e., each of the players wants to maximize its payoff.

Such rational behavior is based on information the player can obtain associated with the network parameters, i.e. the communication environment and scenario. In this paper using a game-theoretical approach we consider a dilemma

that arises in bandwidth scanning problems. On one hand, using tools that allow for the estimation of network parameters allows a player to act rationally to maximize its payoff. On the other hand, such engagement assumes an extra scanning cost. Meanwhile, not engaging such tools by a player allows to reduce involved expenses, although such non-action might also cause a reduction in detection probability since in such a case the player has to design strategy based on the principle of insufficient reasons (also called principle of indifference). In this paper we solve this dilemma via finding equilibrium strategies in the closed form of a non-zero sum stochastic game.

2 Non-Zero Sum Stochastic Game

In this section, we describe our model, which involves a scenario where a *primary user* owns a collection of frequency bands. An adversary, called the *Invader*, attempts to “sneak” usage from some of these bands. The primary user can scan bands to detect such malicious activity. Motivated by such functionality we call the primary user the *Scanner*. All actions (scanning by Scanner and sneaking by Invader) are performed in discrete time slots $1, 2, \dots, \infty$, and the game continues until the Invader is detected. If the Invader is detected, it is eliminated and the game is over.

We assume that at each time slot Scanner and Invader can choose between rational and indifferent modes to scan and sneak, respectively. Denote these mode by \mathbb{R} and \mathbb{I} , respectively.

- (a) In rational mode, the player acts based on information obtained by using special tools to estimate the bands’ parameters so as to adjust its efforts for optimization of the corresponding output.
- (b) In indifferent mode, the player does not engage such tools due to the expenses involved. This leads to a lack of information about the network parameters, and makes the player implement a strategy based on other principles, such as the principle of insufficient reasons (principle of indifference) [13]. In this case, the player spreads its efforts equally across the bands.

Let C_S and C_{S0} (C_I and C_{I0}) be costs for employing rational and indifferent modes by Scanner (Invader). Of course,

$$C_S > C_{S0} \text{ and } C_I > C_{I0}. \quad (1)$$

Let α^1 and α^0 be probabilities to detect the Invader if both players implement rational modes or indifferent modes, respectively. Let α be the probability to detect the Invader if the players implement different modes. To avoid bulkiness in the formulas following the basic example of one-shot bandwidth scanning model given in Sect. 3 below we assume that

$$0 < \alpha \leq \alpha^1 < \alpha^0 < 1. \quad (2)$$

Other possible relationships between detection probabilities could be investigated following the suggested in this paper approach.

At each time slot, the instantaneous payoff to Invader is the difference between the probability of its non-detection and cost for the employed sneaking mode. Similarly, at each time slot the instantaneous payoff to the Scanner is the difference between probability of Invader’s detection and cost for employed scanning mode. If the Invader is not detected then the game moves to the next time slot and is played recursively with discount factor δ , which can be considered as a measure of urgency for sneaking usage of the radio spectrum. We describe this non-zero sum stochastic game via bimatrix form (Γ_S, Γ_I) (see, (3) below), where Γ_S and Γ_I are Scanner and Invader components of this game and each entry corresponds to an action pair of the Scanner and Invader. For example, if both the Scanner and Invader use their mode \mathbb{R} , then the Invader is not detected with probability $\bar{\alpha}^1$ where $\bar{\xi} \triangleq 1 - \xi$. In this case the instantaneous payoff to the Scanner is zero and to the Invader is one, and the game will be replayed with discount rate δ . Invader is detected with probability α^1 . In this case, the instantaneous payoff to Scanner is one and to the Invader it is zero, and the game is over.

Thus, this non-zero-sum stochastic game can be presented in recursive form as follows:

$$(\Gamma_S, \Gamma_I) = (M_S(\Gamma_S), M_I(\Gamma_I)), \tag{3}$$

where

$$M_S(\Gamma_S) \triangleq \begin{matrix} & \mathbb{R} & \mathbb{I} \\ \mathbb{R} & \left(\alpha^1 + \bar{\alpha}^1 \delta \Gamma_S - C_S & \alpha + \bar{\alpha} \delta \Gamma_S - C_S \right) \\ \mathbb{I} & \left(\alpha + \bar{\alpha} \delta \Gamma_S - C_{S0} & \alpha^0 + \bar{\alpha}^0 \delta \Gamma_S - C_{S0} \right) \end{matrix}, \tag{4}$$

$$M_I(\Gamma_I) \triangleq \begin{matrix} & \mathbb{R} & \mathbb{I} \\ \mathbb{R} & \left(\bar{\alpha}^1 + \bar{\alpha}^1 \delta \Gamma_I - C_I & \bar{\alpha} + \bar{\alpha} \delta \Gamma_I - C_{I0} \right) \\ \mathbb{I} & \left(\bar{\alpha} + \bar{\alpha} \delta \Gamma_I - C_I & \bar{\alpha}^0 + \bar{\alpha}^0 \delta \Gamma_I - C_{I0} \right) \end{matrix}. \tag{5}$$

We solve this game in *stationary strategies*, i.e., strategies that are independent of history and the current time. Let $\mathbf{x}^T = (x, \bar{x})$ be the stationary (*mixed*) strategy for Scanner assigning the probabilities x and \bar{x} to using actions \mathbb{R} and \mathbb{I} .

Let $\mathbf{y}^T = (y, \bar{y})$ be the stationary (*mixed*) strategy for Invader assigning the probabilities y and \bar{y} to using actions \mathbb{R} and \mathbb{I} , respectively.

Thus, the Scanner and Invader have the same set of feasible stationary (mixed) strategies denoted by \mathcal{S} .

By (3)–(5), stationary equilibrium strategies \mathbf{x} and \mathbf{y} of Scanner and Invader, respectively, have to be functions of the corresponding accumulated payoffs, i.e., $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}(\mathbf{v}), \mathbf{y}(\mathbf{v}))$ where $\mathbf{v} = (v_S, v_I)$ with v_S and v_I are the accumulated payoffs corresponding these equilibrium strategies. By (3)–(5), $(\mathbf{x}(\mathbf{v}), \mathbf{y}(\mathbf{v}))$ is a stationary equilibrium if and only if the following equations hold [18]:

$$v_S = (\mathbf{x}(\mathbf{v}))^T M_S(v_S) \mathbf{y}(\mathbf{v}), \quad (6)$$

$$v_I = (\mathbf{x}(\mathbf{v}))^T M_I(v_I) \mathbf{y}(\mathbf{v}), \quad (7)$$

where $\mathbf{v} = (v_S, v_I)$ and

$$\mathbf{x}(\mathbf{v}) = \operatorname{argmax}_{\mathbf{x} \in \mathcal{S}} \mathbf{x}^T M_S(v_S) \mathbf{y}(\mathbf{v}), \quad (8)$$

$$\mathbf{y}(\mathbf{v}) = \operatorname{argmax}_{\mathbf{y} \in \mathcal{S}} (\mathbf{x}(\mathbf{v}))^T M_I(v_I) \mathbf{y}. \quad (9)$$

Thus, by (8), for fixed \mathbf{v} and $\mathbf{y}(\mathbf{v})$, $\mathbf{x}(\mathbf{v})$ is the best response to $\mathbf{y}(\mathbf{v})$. By (9), for fixed \mathbf{v} and $\mathbf{x}(\mathbf{v})$, $\mathbf{y}(\mathbf{v})$ is the best response to $\mathbf{x}(\mathbf{v})$. Also, v_A and v_I are such that (6) and (7) also have to hold. Thus, (6)–(9) is the non-zero-sum, and for the Shapley-Bellman equation for the zero-sum game (see, [18]). In Sect. 4 below we solve in closed form, these Shapley(-Bellman) equations.

Finally, we note that in the literature the stochastic game approach has been employed to model different network security problems [4, 8, 9, 17, 23].

3 A Basic Example of Rational and Indifferent Behaviour

In this section, we give a basic model as an example of rational and indifferent behaviour in bandwidth scanning/sneaking. The model involves a scenario where a *primary user* (Scanner) owns n frequency bands $1, 2, \dots, n$. Invader will attempt to “sneak” usage on *only* one of these bands. By assumption, the Scanner can only scan a single band at a time to detect such malicious activity. We assume that the Invader will be detected with probability γ_i , $\gamma_i \in (0, 1)$, if it sneaks in band i and Scanner scans that band. If Scanner does not scan the band that Invader is using, then the Invader sneaks safely, i.e., its detection probability is zero. We assume that the payoff to the Scanner is one if the Invader is detected, and it is zero otherwise. Note that in optimization framework such kind of problems relates to search theory founded by B. Koopman [14].

Let a (mixed) strategy for the Scanner be $\mathbf{p} = (p_1, \dots, p_n)$, where p_i is the probability that it scans band i . So, $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$, $i = 1, \dots, n$.

Let a (mixed) strategy for the Invader be $\mathbf{q} = (q_1, \dots, q_n)$, where q_i is the probability that it sneaks in band i . Thus, $\sum_{i=1}^n q_i = 1$ and $q_i \geq 0$, $i = 1, \dots, n$. Then, the expected payoff to the Scanner, which reflects the detection probability of the Invader, for the players employing strategies \mathbf{p} and \mathbf{q} , is given as follows:

$$V(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n \gamma_i p_i q_i. \quad (10)$$

3.1 Rational Players

Next we consider when both players are rational, i.e., they know the network parameters. Based on this information the Scanner wants to maximize detection probability of the Invader, i.e., to maximize $V(\mathbf{p}, \mathbf{q})$ on \mathbf{p} for each fixed \mathbf{q} .

Meanwhile, the Invader wants to minimize such probability. Thus, this is a zero sum game [18]. We look for (Nash) equilibrium strategies. Recall that (\mathbf{p}, \mathbf{q}) is an equilibrium in the zero-sum game if and only if the following inequalities hold for all feasible $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$:

$$V(\tilde{\mathbf{p}}, \mathbf{q}) \leq V(\mathbf{p}, \mathbf{q}) \leq V(\mathbf{p}, \tilde{\mathbf{q}}) \tag{11}$$

This is a zero sum game with a diagonal payoff matrix and its equilibrium strategies are given in closed form as follows [7]:

$$p_i^1 = q_i^1 = \frac{1/\gamma_i}{\sum_{j=1}^n (1/\gamma_j)}, i = 1, \dots, n \tag{12}$$

with the equilibrium payoff to the Scanner, which reflects the detection probability, being equal to

$$\alpha^1 \triangleq V(\mathbf{p}^1, \mathbf{q}^1) = \frac{1}{\sum_{j=1}^n (1/\gamma_j)}. \tag{13}$$

3.2 Strategies Based on Principle of Indifference

If a player does not have tools (for example, to estimate band’s parameters reflected by detection probabilities) to adjust its efforts, or it does not engage such tools due to the expenses involved, the player must design strategy based on other principles, such as the principle of insufficient reasons (principle of indifference) [13]. This principle might be applied in absence of information regarding the network’s bands. In this case, the player would spread its effort equally among the bands.

Thus, based on the principle of insufficient reasons, the Invader implements strategy \mathbf{q}^0 which uniformly distributes sneaking efforts over bands, i.e.,

$$\mathbf{q}^0 \triangleq (1/n, \dots, 1/n). \tag{14}$$

Similarly, Scanner, based on the principle of insufficient reasons, implements strategy \mathbf{p}^0 , which uniformly distributes scanning efforts over bands, i.e.,

$$\mathbf{p}^0 \triangleq (1/n, \dots, 1/n). \tag{15}$$

3.3 Relations Between Detection Probabilities

In the following lemma we establish relationships between the detection probabilities for different combinations of player’s behavior.

Lemma 1. (a) *If both players implement strategies based on principle of indifference then the detection probability is equal to*

$$\alpha^0 \triangleq V(\mathbf{p}^0, \mathbf{q}^0) = \sum_{i=1}^n \gamma_i/n^2. \tag{16}$$

(b) *If only one player implements strategy based on principle of indifference then the detection probability is equal to*

$$\alpha = V(\mathbf{p}^1, \mathbf{q}^0) = V(\mathbf{p}^0, \mathbf{q}^1) = \frac{1}{\sum_{i=1}^n (1/\gamma_i)}. \tag{17}$$

(c) *Between detection probabilities the following relations hold:*

$$0 < \alpha = \alpha^1 < \alpha^0 < 1. \tag{18}$$

Note that (18) is a particular case of assumption (2).

The proof of Lemma 1 can be found in Appendix A.1.

4 Solution of Non-Zero Sum Stochastic Game

In this section we design equilibrium strategies in closed form. First we derive conditions when equilibrium is in pure strategies. In the remaining cases, we derive equilibrium in mixed strategies and prove its uniqueness.

4.1 Equilibrium in Pure Strategies

In this section we establish conditions for the equilibrium to be in pure strategies.

A Pair of Actions (\mathbb{R}, \mathbb{R}) is Equilibrium: Let us consider the situation where a pair of actions (\mathbb{R}, \mathbb{R}) is equilibrium. In other words, when $((1, 0), (1, 0))$ is equilibrium.

Theorem 1. *In the stochastic game Γ , a pair of actions (\mathbb{R}, \mathbb{R}) is an equilibrium with corresponding accumulated payoffs (v_S, v_I) if and only if*

$$v_S = \frac{\alpha^1 - C_S}{1 - \overline{\alpha^1}\delta} \text{ and } v_I = \frac{\overline{\alpha^1} - C_I}{1 - \overline{\alpha^1}\delta} \tag{19}$$

and

$$\frac{1 - \overline{\alpha}\delta}{1 - \overline{\alpha^1}\delta}(\alpha_1 - C_S) > \alpha - C_{S0} \text{ and } \frac{1 - \overline{\alpha}\delta}{1 - \overline{\alpha^1}\delta}(\overline{\alpha_1} - C_I) > \overline{\alpha} - C_{I0}. \tag{20}$$

The proof can be found in Appendix A.2.

A Pair of Actions (\mathbb{I}, \mathbb{I}) is Equilibrium: Let us examine a situation where a pair of actions (\mathbb{I}, \mathbb{I}) is an equilibrium. In other words, when $((0, 1), (0, 1))$ is equilibrium.

Theorem 2. *In the stochastic game Γ , a pair of actions (\mathbb{I}, \mathbb{I}) is an equilibrium with corresponding accumulated payoffs (v_S, v_I) if and only if*

$$v_S = \frac{\alpha^0 - C_{S0}}{1 - \overline{\alpha^0}\delta} \text{ and } v_I = \frac{\overline{\alpha^0} - C_{I0}}{1 - \overline{\alpha^0}\delta} \tag{21}$$

and

$$\frac{1 - \overline{\alpha}\delta}{1 - \overline{\alpha^0}\delta}(\alpha_0 - C_{S0}) > \alpha - C_S \text{ and } \frac{1 - \overline{\alpha}\delta}{1 - \overline{\alpha^0}\delta}(\overline{\alpha_0} - C_{I0}) > \overline{\alpha} - C_I. \tag{22}$$

The proof can be found in Appendix A.3.

A Pair of Actions (\mathbb{R}, \mathbb{I}) is Equilibrium: Let us consider situation where a pair of actions (\mathbb{R}, \mathbb{I}) is an equilibrium. In other words, when $((1, 0), (0, 1))$ is an equilibrium.

Theorem 3. *In the stochastic game Γ , a pair of actions (\mathbb{R}, \mathbb{I}) is an equilibrium with corresponding accumulated payoffs (v_S, v_I) if and only if*

$$v_S = \frac{\alpha - C_S}{1 - \bar{\alpha}\delta} \text{ and } v_I = \frac{\bar{\alpha} - C_{I0}}{1 - \bar{\alpha}\delta} \tag{23}$$

and

$$\frac{1 - \bar{\alpha}^0\delta}{1 - \bar{\alpha}\delta}(\alpha - C_S) > \alpha^0 - C_{S0} \text{ and } \frac{1 - \bar{\alpha}^1\delta}{1 - \bar{\alpha}\delta}(\bar{\alpha}_0 - C_{I0}) > \bar{\alpha}^1 - C_I. \tag{24}$$

The proof can be found in Appendix A.4.

A Pair of Actions (\mathbb{I}, \mathbb{R}) is an Equilibrium: Let us consider a situation where a pair of actions (\mathbb{I}, \mathbb{R}) is an equilibrium. In other words, when $((0, 1), (1, 0))$ is equilibrium.

Theorem 4. *In the stochastic game Γ , a pair of actions (\mathbb{I}, \mathbb{R}) is an equilibrium with corresponding accumulated payoffs (v_S, v_I) if and only if*

$$v_S = \frac{\alpha - C_{S0}}{1 - \bar{\alpha}\delta} \text{ and } v_I = \frac{\bar{\alpha} - C_I}{1 - \bar{\alpha}\delta} \tag{25}$$

and

$$\frac{1 - \bar{\alpha}^1\delta}{1 - \bar{\alpha}\delta}(\alpha - C_{S0}) > \alpha^1 - C_S \text{ and } \frac{1 - \bar{\alpha}^0\delta}{1 - \bar{\alpha}\delta}(\bar{\alpha} - C_I) > \bar{\alpha}^0 - C_{I0}. \tag{26}$$

The proof can be found in Appendix A.5.

Uniqueness of Equilibrium in Pure Strategies: Let us prove that only one equilibrium in pure strategies could exist.

Theorem 5. (a) *None of a pair of conditions (20), (22), (24) and (26) can hold simultaneously.*

(b) *Only one equilibrium in pure strategies could exist.*

The proof can be found in Appendix A.6.

4.2 Equilibrium in Mixed Strategies

In this section we consider the situation where the equilibrium (\mathbf{x}, \mathbf{y}) is in mixed strategies, i.e., $x \in (0, 1)$ and $y \in (0, 1)$.

Theorem 6. *In the stochastic game Γ , the equilibrium is unique. Moreover,*

- (a) *if one of four conditions (20), (22), (24) and (26) holds, then equilibrium is in pure strategies, and it is given by Theorem 1–Theorem 4, respectively;*
- (b) *if none of the four conditions (20), (22), (24) and (26) holds then equilibrium is in mixed strategies. Moreover, it is equal to $((x, \bar{x}), (y, \bar{y}))$ with corresponding accumulated payoffs (v_S, v_I) , where*

$$v_S = \frac{C_S - \alpha}{\alpha^1 - \alpha} + \frac{C_{S0} - \alpha^0}{\alpha^0 - \alpha} \quad \text{and} \quad v_I = \frac{C_I - \bar{\alpha}}{\alpha^1 - \alpha} + \frac{C_{I0} - \bar{\alpha}^0}{\alpha^0 - \alpha} \tag{27}$$

$$\frac{\bar{\alpha}\delta - 1}{\alpha^1 - \alpha} + \frac{\bar{\alpha}^0\delta - 1}{\alpha^0 - \alpha}$$

and

$$x = \frac{C_I + \alpha + (1 - \bar{\alpha}\delta)v_I}{(\alpha - \alpha^1)(1 + \delta v_I)} \quad \text{and} \quad y = \frac{C_S - \alpha + (1 - \bar{\alpha}\delta)v_S}{(\alpha^1 - \alpha)(1 - \delta v_S)}. \tag{28}$$

The proof can be found in Appendix A.7.

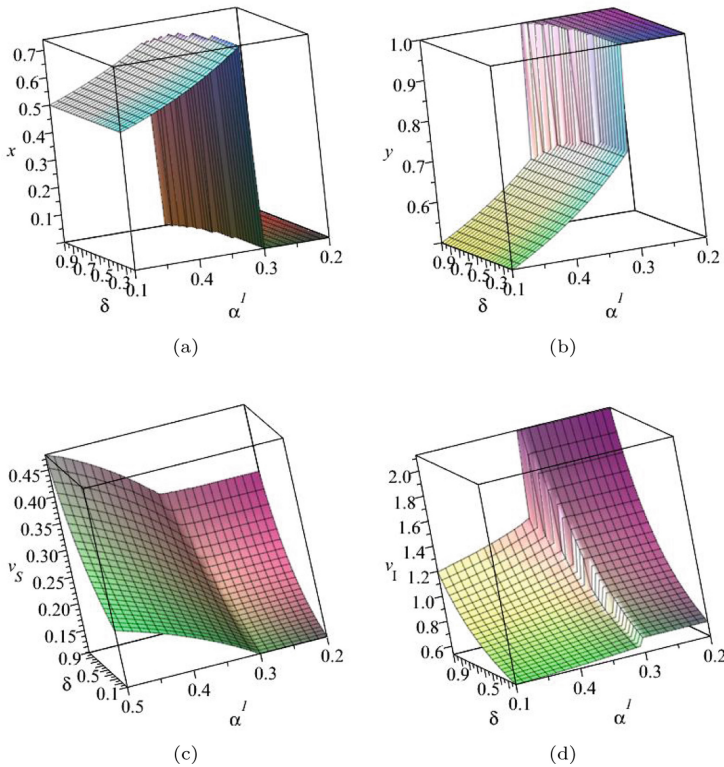


Fig. 1. (a) Probability x , (b) probability y (c) payoff to Scanner and (d) payoff to Invader as functions on detection probability α^1 and discount factor δ .

5 Discussion of the Results

Let us illustrate the obtained results on a network with scanning costs $C_{S0} = 0.1$ and $C_S = 0.2$, sneaking costs $C_{I0} = 0.1$ and $C_I = 0.2$, detection probability $\alpha^0 = 0.5$ when both players implement indifferent mode and detection probability $\alpha = 0.2$ when only one of the players implements indifferent mode. Figure 1 illustrates that an increase of discount factor leads to an increase in payoffs of both players. An increase in detection probability α^1 when both players implement a rational mode leads to an increase in payoff to Scanner and a decrease in payoff to Invader. For small probability α^1 , the players follow pure strategies implementing opposite modes. For large probability α^1 the players implement mixed strategies.

6 Conclusions

A dilemma, arising in bandwidth scanning problems, that occurs because the agents implement strategies designed based on either a rational principle or the principle of insufficient reasons, has been modeled as a non-zero sum stochastic game between two players (Scanner and Invader). The equilibrium for this stochastic game has been found in closed form in stationary strategies via solving the corresponding Shapley(-Bellman) equations, and its uniqueness has been proven. The proven uniqueness of equilibrium reflects stability of the designed trade-off selection between scanning strategies designed based on a rational principle and the principle of insufficient reasons.

Acknowledgement. This work was supported in part by the U.S. National Science Foundation under grants CNS-1909186 and ECCS-2128451.

A Appendix

A.1 Proof of Lemma 1

First, note that (12), (14) and (15) imply (a) and (b).

By symmetry, function $\sum_{i=1}^n \gamma_i \sum_{i=1}^n (1/\gamma_i)$ achieves its minimum at such $\gamma_i \in [0, 1]$, $i = 1, \dots, n$ that $\gamma_1 = \dots = \gamma_n$. Meanwhile,

$$\sum_{i=1}^n \gamma_i \sum_{i=1}^n (1/\gamma_i) = n^2 \text{ for } \gamma_1 = \dots = \gamma_n. \tag{29}$$

This jointly with (13) and (16), imply that

$$\alpha^1 < \alpha^0. \tag{30}$$

Finally, (b) and (30) imply (18), and the result follows. ■

A.2 Proof of Theorem 1

By (6)–(9), we have that (\mathbb{R}, \mathbb{R}) is equilibrium with v_S and v_I as the corresponding accumulated payoffs if and only if the following relations hold:

$$\alpha^1 + \bar{\alpha}^1 \delta v_S - C_S = v_S, \tag{31}$$

$$\bar{\alpha}^1 + \bar{\alpha}^1 \delta v_I - C_I = v_I \tag{32}$$

with

$$\alpha^1 + \bar{\alpha}^1 \delta v_S - C_S > \alpha + \bar{\alpha} \delta v_S - C_{S0}, \tag{33}$$

$$\bar{\alpha}^1 + \bar{\alpha}^1 \delta v_I - C_I > \bar{\alpha} + \bar{\alpha} \delta v_I - C_{I0}. \tag{34}$$

Solving (31) and (32) by v_S and v_I , respectively, imply (19). Substituting v_S and v_I given by (19) into (33) and (34) imply (20). ■

A.3 Proof of Theorem 2

By (6)–(9), we have that (\mathbb{I}, \mathbb{I}) with v_S and v_I as the corresponding accumulated payoffs if and only if the following relations hold:

$$\alpha^0 + \bar{\alpha}^0 \delta v_S - C_{S0} = v_S, \tag{35}$$

$$\bar{\alpha}^0 + \bar{\alpha}^0 \delta v_I - C_{I0} = v_I \tag{36}$$

with

$$\alpha^0 + \bar{\alpha}^0 \delta v_S - C_{S0} > \alpha + \bar{\alpha} \delta v_S - C_S, \tag{37}$$

$$\bar{\alpha}^0 + \bar{\alpha}^0 \delta v_I - C_{I0} > \bar{\alpha} + \bar{\alpha} \delta v_I - C_I. \tag{38}$$

Solving (35) and (36) by v_S and v_I , respectively, imply (21). Substituting v_S and v_I given by (21) into (37) and (38) imply (22). ■

A.4 Proof of Theorem 3

By (6)–(9), we have that (\mathbb{R}, \mathbb{I}) is an equilibrium with v_S and v_I as the corresponding accumulated payoffs if and only if the following relations hold:

$$\alpha + \bar{\alpha} \delta v_S - C_S = v_S, \tag{39}$$

$$\bar{\alpha} + \bar{\alpha} \delta v_I - C_{I0} = v_I \tag{40}$$

with

$$\alpha + \bar{\alpha} \delta v_S - C_S > \alpha^0 + \bar{\alpha}^0 \delta v_S - C_{S0}, \tag{41}$$

$$\bar{\alpha} + \bar{\alpha} \delta v_I - C_{I0} > \bar{\alpha}^1 + \bar{\alpha}^1 \delta v_I - C_I. \tag{42}$$

Solving (39) and (40) by v_S and v_I , respectively, imply (23). Substituting v_S and v_I given by (23) into (41) and (42) imply (24). ■

A.5 Proof of Theorem 4

By (6)–(9), we have that (\mathbb{I}, \mathbb{R}) is an equilibrium with v_S and v_I as the corresponding accumulated payoffs if and only if the following relations hold:

$$\alpha + \bar{\alpha}\delta v_S - C_{S0} = v_S, \tag{43}$$

$$\bar{\alpha} + \bar{\alpha}\delta v_I - C_I = v_I \tag{44}$$

with

$$\alpha + \bar{\alpha}\delta v_S - C_{S0} > \alpha^1 + \bar{\alpha}^1\delta v_S - C_S, \tag{45}$$

$$\bar{\alpha} + \bar{\alpha}\delta v_I - C_I > \bar{\alpha}^0 + \bar{\alpha}^0\delta v_I - C_{I0} \tag{46}$$

Straightforward calculation based on (43)–(46) implies the result. ■

A.6 Proof of Theorem 5

(b) follows from (a) and Theorem 1–Theorem 4.

Let us now prove (a). First note that by (2) we have that

$$\bar{\alpha} \geq \bar{\alpha}^1 > \bar{\alpha}^0. \tag{47}$$

Assume that (20) and (22) hold simultaneously. Then, (47) and the second of inequalities (20) imply

$$\bar{\alpha} - C_I > \frac{1 - \bar{\alpha}^1\delta}{1 - \bar{\alpha}\delta}(\bar{\alpha} - C_{I0}). \tag{48}$$

Combining this inequality with the second of inequalities (20) imply

$$\frac{1 - \bar{\alpha}\delta}{1 - \bar{\alpha}^0\delta}(\bar{\alpha}^0 - C_{I0}) > \bar{\alpha} - C_I > \frac{1 - \bar{\alpha}^1\delta}{1 - \bar{\alpha}\delta}(\bar{\alpha} - C_{I0}). \tag{49}$$

Meanwhile, by (47),

$$(1 - \bar{\alpha}\delta)^2 < (1 - \bar{\alpha}^0\delta)1 - \bar{\alpha}^1\delta \tag{50}$$

and

$$\bar{\alpha}^0 - C_{I0} < \bar{\alpha} - C_{I0} \tag{51}$$

By (50) and (51), we have that

$$\frac{1 - \bar{\alpha}\delta}{1 - \bar{\alpha}^0\delta}(\bar{\alpha}^0 - C_{I0}) < \frac{1 - \bar{\alpha}^1\delta}{1 - \bar{\alpha}\delta}(\bar{\alpha} - C_{I0}). \tag{52}$$

This contradicts to (49). Thus, conditions (20) and (22) cannot hold simultaneously.

The remaining pairs of conditions can be checked similarly. ■

A.7 Proof of Theorem 6

First note that there exists at least one equilibrium in stochastic game Γ , since Γ is a recursive game with discount factor $\delta \in (0, 1)$ [18].

By proofs of Theorem 1–Theorem 4, it follows that if equilibrium is in pure strategies, then it is given by Theorem 1–Theorem 4, and one of four conditions (20), (22), (24) and (26) have to hold. Moreover, by Theorem 5, only one of these four conditions can hold, and (a) follows.

Thus, if none of the four conditions (20), (22), (24) and (26) holds then the equilibrium is in mixed strategies. By (6)–(9), we have that $((x, \bar{x}), (y, \bar{y}))$ with $x \in (0, 1)$ and $y \in (0, 1)$ is an equilibrium with v_S and v_I as the corresponding accumulated payoffs if and only if the following relations hold:

$$(\alpha^1 + \bar{\alpha}^1 \delta v_S - C_S)y + (\alpha + \bar{\alpha} \delta v_S - C_S)\bar{y} = v_S, \tag{53}$$

$$(\alpha + \bar{\alpha} \delta v_S - C_{S0})y + (\alpha^0 + \bar{\alpha}^0 \delta v_S - C_{S0})\bar{y} = v_S \tag{54}$$

and

$$(\bar{\alpha}^1 + \bar{\alpha}^1 \delta v_I - C_I)x + (\bar{\alpha} + \bar{\alpha} \delta v_I - C_I)\bar{x} = v_I, \tag{55}$$

$$(\bar{\alpha} + \bar{\alpha} \delta v_I - C_{I0})x + (\bar{\alpha}^0 + \bar{\alpha}^0 \delta v_I - C_{I0})\bar{x} = v_I. \tag{56}$$

Solving (53) by y implies the second of two Eqs. (28). Solving (54) by y , we have that

$$y = \frac{C_{S0} - \alpha^0 + (1 - \bar{\alpha}^0 \delta)v_S}{(\alpha - \alpha^0)(1 - \delta v_S)}. \tag{57}$$

Substituting this y into the second of two Eqs. (28) and solving by v_S implies the first of two Eqs. (27).

Solving (55) by x implies the first of (28). Solving (56) by x implies

$$x = \frac{C_{I0} - \bar{\alpha}^0 + (1 - \bar{\alpha}^0 \delta)v_I}{(\alpha^0 - \alpha)(1 + \delta v_I)}. \tag{58}$$

Substituting this x into the first of two Eqs. (28) and solving by v_I implies the second of two Eqs. (27). ■

References




1. Anindya, I.C., Kantarcioglu, M.: Adversarial anomaly detection using centroid-based clustering. In: IEEE International Conference on Information Reuse and Integration (IRI), pp. 1–8 (2018)
2. Baston, V., Garnaev, A.: A search game with a protector. *Nav. Res. Logist.* **47**, 85–96 (2000)
3. Belavagi, M.C., Muniyal, B.: Game theoretic approach towards intrusion detection. In: Proceedings of the International Conference on Inventive Computation Technologies (ICICT) (2016)

4. Calinescu, G., Kapoor, S., Qiao, K., Shin, J.: Stochastic strategic routing reduces attack effects. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 1–5 (2011)
5. Comaniciu, C., Mandayam, N.B., Poor, H.V.: A Wireless Networks Multiuser Detection in Cross-Layer Design. Springer, New York (2005). <https://doi.org/10.1007/0-387-27750-1>
6. Dambreville, F., Le Cadre, J.P.: Detection of a Markovian target with optimization of the search efforts under generalized linear constraints. *Nav. Res. Logist.* **49**, 117–142 (2002)
7. Garnaev, A.: A remark on a helicopter and submarine game. *Nav. Res. Logist.* **40**, 745–753 (1993)
8. Garnaev, A., Baykal-Gursoy, M., Poor, H.V.: A game theoretic analysis of secret and reliable communication with active and passive adversarial modes. *IEEE Trans. Wireless Commun.* **15**, 2155–2163 (2016)
9. Garnaev, A., Hayel, Y., Altman, E., Avrachenkov, K.: Jamming game in a dynamic slotted ALOHA network. In: Jain, R., Kannan, R. (eds.) *GameNets 2011*. LNCS, vol. 75, pp. 429–443. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30373-9_30
10. Garnaev, A., Trappe, W., Kung, C.T.: Dependence of optimal monitoring strategy on the application to be protected. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 1054–1059 (2012)
11. Garnaev, A., Trappe, W., Kung, C.-T.: Optimizing scanning strategies: selecting scanning bandwidth in adversarial RF environments. In: Proceedings of the 8th International Conference on Cognitive Radio Oriented Wireless Networks (Crowncom), pp. 148–153 (2013)
12. Guan, S., Wang, J., Jiang, C., Tong, J., Ren, Y.: Intrusion detection for wireless sensor networks: a multi-criteria game approach. In: Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2018)
13. Jaynes, E.T.: *Probability Theory: The Logic of Science*. Cambridge University Press, New York (2003)
14. Koopman, B.: *Search and Screening. General Principles with Historical Applications*. Pergamon Press (1980)
15. Liu, S., Chen, Y., Trappe, W., Greenstein, L.J.: ALDO: an anomaly detection framework for dynamic spectrum access networks. In: Proceedings of the IEEE International Conference on Computer (INFOCOM), pp. 675–683 (2009)
16. Ma, Y., Cao, H., Ma, J.: The intrusion detection method based on game theory in wireless sensor network. In: Proceedings of the First IEEE International Conference on Ubi-Media Computing (2008)
17. Nguyen, K.C., Alpcan, T., Basar, T.: Stochastic games for security in networks with interdependent nodes. In: Proceedings of the International Conference on Game Theory for Network (GameNets), pp. 697–703 (2009)
18. Owen, G.: *Game Theory*. Academic Press, New York (1982)
19. Poongothai, T., Jayarajan, K.: A noncooperative game approach for intrusion detection in mobile adhoc networks. In: Proceedings of the International Conference on Computing, Communication and Networking (2008)
20. Sauder, D.W., Geraniotis, E.: Signal detection games with power constraints. *IEEE Trans. Inf. Theory* **40**, 795–807 (1994)
21. Shen, S.: A game-theoretic approach for optimizing intrusion detection strategy in WSNs. In: Proceedings of the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC) (2011)

22. Vamvoudakis, K.G., Hespanha, J.P., Sinopoli, B., Mo, Y.: Adversarial detection as a zero-sum game. In: Proceedings of the 51st IEEE Conference on Decision and Control (CDC), pp. 7133–7138 (2012)
23. Wang, B., Wu, Y., Liu, K.J.R., Clancy, T.C.: An anti-jamming stochastic game for cognitive radio networks. *IEEE J. Sel. Areas Commun.* **29**, 877–889 (2011)
24. Wang, X., Feng, R., Wu, Y., Che, S., Ren, Y.: A game theoretic malicious nodes detection model in MANETs. In: Proceedings of the IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS), pp. 1–6 (2012)



Port Capacity Leasing Games at Internet Exchange Points

Md Ibrahim Ibne Alam^(✉) , Elliot Anshelevich , and Koushik Kar 

Rensselaer Polytechnic Institute, Troy, NY 12180, USA
alam2@rpi.edu, eanshel@cs.rpi.edu, koushik@ecse.rpi.edu

Abstract. Internet Service Providers (ISPs) lease ports at a public switch in an Internet Exchange Point (IXP) to exchange traffic efficiently with other ISPs present at the IXP. The price paid to lease a port depends on the port capacity, which also impacts the Quality of Service (QoS) experienced by the ISP's traffic exchanged through the IXP switch. In this paper, we analyze the leasing of port capacities at an IXP as a non-cooperative game between the ISPs, and analyze the efficiency at equilibrium as compared to the social optimum. We show that when the IXP switch capacity is not changed in response to the port capacities purchased, there is dominant strategy for each ISP that attains a Price of Anarchy (PoA) of at most 2. If the IXP switch capacity is varied to “match” the aggregate port capacity leased by the ISPs, then bad equilibria can exist. However, under certain reasonable assumptions, the PoA is still guaranteed to be within 2. Simulation studies demonstrate the effect of the per-unit leasing price and switch delay functions on the equilibrium performance; in all scenarios simulated, the social cost at equilibrium was found to be very close to the optimum.

Keywords: Internet exchange point · Internet service provider · Public peering · Port capacity leasing

1 Introduction

Internet eXchange Points (IXPs) are data centers with network switches through which Internet Service Providers (ISPs) exchange traffic, mostly through peering relationships [1, 8]. In spite of falling transit costs, peering between ISPs has been increasing, resulting in flattening of the Internet [1, 7, 12]. It is estimated that almost 80% of the IP addresses in the world can be reached via public peering, and 20% of all the traffic go through IXPs [4]. Exchange at IXPs typically improves traffic Quality of Service (QoS) due to lower delays and losses associated with shorter Internet paths, and offers reachability to a large number of other ISPs that are present at that IXP.

For peering at the public switch at the IXP, ISPs typically pay a price that depends on the port size (i.e., capacity of the port leased by the ISP) [13]. The larger the port capacity leased, the larger the price paid; however, a larger

port capacity also allows faster transfer of traffic. The QoS experienced by the traffic sent by an ISP, however, also depends on the port capacities purchased by other ISPs. This results in the port-capacity leasing decisions of the ISPs being dependent on each other, and whether or not the IXP switch speed (capacity) is varied based on the port capacities purchased by the ISPs.

In this paper, we analyze the port capacity leasing problem for ISPs (at a public switch in an IXP) as a non-cooperative game, where each ISP is making the leasing decision selfishly so as to minimize its own cost in sending the traffic through the IXP. This cost comprises of the port capacity leasing price to be paid to the IXP, as well as the congestion costs at each of the queuing points inside the IXP. We analyze two scenarios: (a) the switch capacity is fixed, i.e., remains unaffected by the port capacities leased by the ISPs; (b) the switch capacity is upgraded/downgraded so as to “match” the aggregate port capacity leased by the ISPs. For each scenario, we evaluate the social cost at equilibrium as compared to the optimum social cost, represented as the Price of Anarchy (PoA) or the Price of Stability (PoS). For scenario (a), we show that there exists a dominant strategy for each ISP, which also attains a PoA of 2. For scenario (b), bad equilibria exist; however, under the reasonable assumption that the switch capacity is varied so that the switch is not the congestion bottleneck in the system, there exist good equilibria and PoA is at most 2. Simulations over a range of parameters, using traffic data for US IXPs as estimated based on their actual locations, show the existence of at least one very good equilibria (PoS very close to 1). Alongside, the values of PoS showed a increasing trend with traffic symmetry and the number of ISPs peering at the IXP.

Our work is related to some of the prior work on the effect of IXP pricing on peering relationships [2,3]; however the problem investigated in this paper and its game theoretic model is fundamentally different from that analyzed in these earlier work. In [2,3] the strategy of each ISP involves determining *how much traffic to send through the IXP versus sending it outside of the IXP* (through transit providers). Further, the price charged to each ISP by the IXP depends on the amount of traffic sent by the ISP through the IXP. In contrast, the model in our paper considers determining *how much port capacity an ISP should lease, given a certain amount of traffic to be sent through the IXP*. Given that prices paid by an ISP at an IXP primarily depend on the port capacity [9,15], this model reflects current practice. The tools used in the game theoretic analysis, and the nature of the results derived, also differ quite significantly. In [2,3] it is observed that the choice of the traffic-dependent pricing policy has a significant effect on how much of the IXP traffic flows through the IXP at equilibrium, thereby impacting the PoA values significantly. In this paper we observe that when the traffic (to be exchanged via IXP) is already determined, the worst case social cost values at equilibrium is relatively unimpacted by the port leasing costs, i.e., good PoA and PoS values (close to 1) exist irrespective of the port leasing costs. Finally, while [2,3] draws upon a long line of work on network formation games (e.g., [11,16]) and utilizes the notion of pairwise equilibrium [6,10], our model is not a network formation game, and studies Nash equilibria in the context of the port capacity leasing problem.

2 System Model and Properties

In this section, we detail our game theoretic model of the port capacity leasing problem, and argue the existence of the Nash equilibrium of the game.

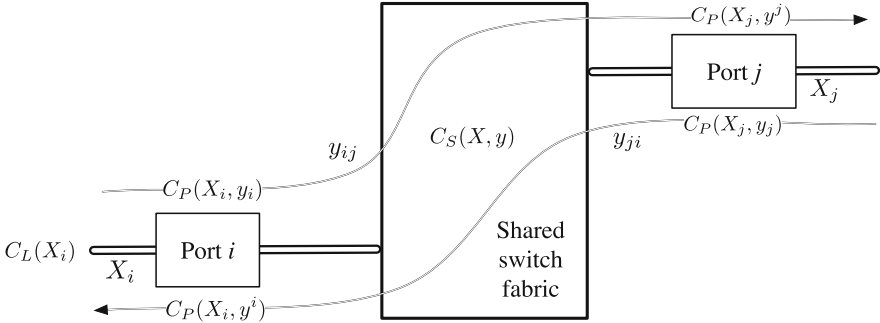


Fig. 1. Schematic of the system model showing different costs faced by agent (ISP) i .

Consider a public (shared) switch at an IXP, which is utilized by N ISPs to exchange traffic between themselves. The amount of traffic that any ISP (say i) has for the other ISP (say j) is directional and given (fixed), and denoted by y_{ij} . We use $y_i = \sum_j y_{ij}$ ($y^i = \sum_j y_{ji}$) to denote the total traffic sent from (to) ISP i to (from) all other ISPs. Let X_i denote the port capacity leased by any ISP i (ISP i 's strategy), which provides a connection bandwidth of X_i (in both directions) between ISP i 's server equipment and the shared switch fabric. Note that for stability, $X_i \geq \max(y_i, y^i), \forall i$. As illustrated in Fig. 1, the traffic from ISP i to ISP j , y_{ij} , can face congestion at three potential bottlenecks (queues on its path in the IXP) that are impacted by the port capacity choices of the ISPs (i, j and potentially other ISPs): (i) A congestion cost of $C_P(X_i, y_i)$ at the port i which has an aggregate traffic of y_i and served by a link of capacity X_i ; (ii) A congestion cost of $C_P(X_j, y^j)$ at port j which has an aggregate traffic of y^j and served by a link of capacity X_j ; (iii) A congestion cost of $C_S(X_i, \dots, y_i, y^i, \dots, \forall i)$ at the shared switch. Note that $(y_i, y^i, \forall i)$ represents the entire traffic vector that shares the public switch, whose fabric capacity may or may not depend on the vector of port capacities leased, $X = (X_i, \forall i)$. Accordingly, in our analysis in Sect. 3 we consider both scenarios: when $C_S(\cdot)$ is only a function of the total traffic vector $y = (y_i, y^i, \forall i)$, and when it is a function of the leased capacity vector X as well. Finally, ISP i also faces a port leasing cost given by $C_L(X_i)$.

Let X_{-i} denote the set of strategies of all ISPs except i . Let $[0, M]$ represent the range over which any port capacity can be chosen. Then in the port capacity leasing game, given X_{-i} , each ISP i will choose X_i in the range $[\max(y_i, y^i), M]$ so as to selfishly minimize its cost function $C'_i(X_i; X_{-i})$ (continuous in X_i), given below.

$$\begin{aligned}
 C_i(X_i; X_{-i}) &= \sum_j \left[y_{ij} \left(C_P(X_i, y_i) + C_P(X_j, y^j) \right) + y_{ji} \left(C_P(X_j, y_j) + C_P(X_i, y^i) \right) \right] \\
 &\quad + C_L(X_i) + \sum_j (y_{ij} + y_{ji}) C_S(X, y) \\
 &= y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i) + \sum_j \left(y_{ij} C_P(X_j, y^j) + y_{ji} C_P(X_j, y_j) \right) \\
 &\quad + C_L(X_i) + \left(y_i + y^i \right) C_S(X, y). \tag{1}
 \end{aligned}$$

The notion of equilibrium for this game will be of Nash Equilibrium, where no single player (ISP) can improve its cost function (as defined by Eq. 1) by unilaterally changing its own port capacity. We make the following reasonable assumption on the congestion cost functions, which will be used in our analyses throughout the paper.

Assumption 1. $C_i(X_i; X_{-i})$ has a unique minimum in X_i for any given X_{-i} .

The functions $C_P(X_i, y_i)$ and $C_P(X_i, y^i)$ are related to the congestion delay at the ports, and can be expected to be decreasing and strictly convex in X_i ; for example, the M/M/1 delay function is decreasing and strictly convex in the server capacity. In practice, $C_L(X_i)$ is usually defined for a few discrete port capacity choices [9, 15], but can be reasonably approximated as an increasing and convex function of X_i . The function $C_S(X, y)$ is related to the congestion delay at the switch fabric, and its exact nature will depend on how the switch fabric capacity is varied as a function of the port capacity choices. However, as it is a congestion delay function, and the switch fabric capacity is not likely to decrease with increase in port capacity, it is reasonable to assume that it will be convex function of X . Therefore, $C_i(X_i, X_{-i})$ can be expected to be strictly convex in X_i , thereby satisfying Assumption 1. We now establish the existence of equilibria for this game, denoted by X_{eq} .

Proposition 1. *Under Assumption 1, an equilibrium for the port capacity leasing game always exists.*

Proof. The Nash equilibrium of the port capacity leasing game is characterized by the following fixed point equations: $X_i = f_i(X_{-i}), \forall i$, where $f_i(X_{-i}) = \arg \min_{X_i \in [0, M]} C_i(X_i; X_{-i})$. Note that Assumption 1 implies a unique best response strategy, i.e., $\arg \min_{X_i \in [0, M]} C_i(X_i; X_{-i})$ produces a unique value. This implies the function f_i is well defined as well as continuous. Therefore, the conditions of the Brouwer’s fixed point theorem are satisfied, implying that there exists a fixed point of the set of equations $X_i = f_i(X_{-i}), \forall i$. \square

Finally, we define the social optimum (OPT), denoted by X^* , as a port capacity leasing solution that minimizes total cost of all the ISPs, expressed as

$$\begin{aligned}
 C(X) &= \sum_i C_i(X_i; X_{-i}) \\
 &= \sum_i \left[y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i) + \sum_j (y_{ij} C_P(X_j, y^j) + y_{ji} C_P(X_j, y_j)) \right] \\
 &\quad + \sum_i [C_L(X_i) + (y_i + y^i) C_S(X, y)] \\
 &= \sum_i [2(y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i)) + C_L(X_i) + (y_i + y^i) C_S(X, y)]. \quad (2)
 \end{aligned}$$

Multiple Equilibria: In many cases of practical interest, the port leasing game has unique equilibrium (as will be discussed in Sect. 3.1). However there can be cases where multiple equilibria exist; one such example involving 2-ISPs are as follows. Let us assume that the 2 ISPs (say i and j) have $y_{ij} = y_{ji} = 1$ amount of traffic to exchange. Also, $C_P(X_i) = k$ (constant), $C_S(X) = \max(10 - (\sum_i X_i - y), 0)$, and $C_L(X_i) = \log(X_i)$. Then with $C_i(X_i, X_{-i}) = 4k + 2(12 - \sum_i X_i) + \log(X_i), \forall i$, all the combinations of $X_i + X_j = 12$, with $X_i \geq 1$ and $X_j \geq 1$ is an equilibrium solution. For example, $X_i = 8, X_j = 4$ is one such equilibrium, where none of the ISPs can improve their cost by unilaterally changing their leased port capacity.

3 Price of Anarchy Analysis

The Price-of-Anarchy (*PoA*) (Price-of-Stability (*PoS*)) is defined as the ratio of the cost at worst (best) equilibrium to the cost at OPT. We analyze two scenarios, based on whether the switch congestion function C_S remains fixed or varies as a function of the leased port capacity vector, X . The first scenario, in which C_S is independent of X , the switch capacity is given and not varied based on leased port capacity. The second scenario models the case where the switch fabric capacity is provisioned based on leased port capacity. The two scenarios require different analytical treatments, and therefore discussed separately.

3.1 Fixed Switch Capacity

When $C_S(X, y)$ is independent of the strategy vector X (but may depend on the traffic vector y), the port capacity leasing game becomes a potential game, and we obtain the following properties of the equilibrium solution.

Theorem 1. *Under Assumption 1, if $C_S(X, y)$ is independent of X , then:*

- (i) *Each ISP has a dominant strategy, and the port capacity leasing game has a unique equilibrium.*

(ii) $PoA = PoS \leq 2$.

Proof. *i)* When $C_S(X, y)$ is independent of X , any ISP (say i) can only change $y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i) + C_L(X_i)$ part of its cost (Eq. 1) by changing X_i . Hence from Assumption 1, any ISP will always have a dominant strategy (choosing the X_i resulting in minimum cost), and all the ISPs choosing that specific port capacities will result in a unique equilibrium.

ii) Let us define a function $\Phi(X)$ as follows:

$$\Phi(X) = \sum_i [(y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i)) + C_L(X_i) + (y_i + y^i) \cdot C_S], \quad (3)$$

where X denotes the set of strategies (port capacity choices) of all ISPs. Also, let $\Phi(X'_i, X_{-i})$ express the function for the case where ISP i chooses the strategy X'_i , while the other ISPs choose X_{-i} . Then it is straightforward to show the following:

$$\Phi(X_i, X_{-i}) - \Phi(X'_i, X_{-i}) = C_i(X_i, X_{-i}) - C_i(X'_i, X_{-i}). \quad (4)$$

Hence, according to the definition of potential game [16], we have a potential game with a potential function $\Phi(X)$. Also, from potential game analysis [16], if the total cost $C(X) \leq a \cdot \Phi(X)$, for some scalar a , then that game has a $PoS \leq a$. For our current case, $C(X) \leq 2\Phi(X)$, and hence we get $PoS \leq 2$ for this game. Lastly, from part *i)* of this proof we know that there is a unique equilibrium and hence we have $PoA = PoS \leq 2$. \square

3.2 Variable Switch Capacity

The port capacity leasing game in general does not have a potential function when C_S is a function of the port capacities purchased by the ISPs. As shown later in this section, under certain reasonable additional assumptions the PoA can be shown to be small. However, for the general problem (without any additional restrictions) there can be examples which results in very bad PoA ; one such example is discussed next.

Bad PoA example: For this example we assume that C_P is zero; $C_S = 2N - \sum_i X_i$, where N is the total number of ISPs and X_i the port capacities of those ISPs; and $C_L(X_i) = (2 + \epsilon)X_i$, where ϵ is a small positive value. Also, let's assume that $y_i = y^i = 1, \forall i$. If all the ISPs except i decide to buy unit port capacity, i.e., $X_j = 1, \forall j \neq i$, then from Eq. 1 we get:

$$C_i(X) = (2 + \epsilon)X_i + 2(2N - (N - 1) - X_i) = 2N + 2 + \epsilon X_i. \quad (5)$$

So, from the perspective of ISP i , it would want to buy the least possible port capacity, which is $X_i = \max(y_i, y^i)$. Hence, ISP i will end up choosing $X_i = 1$. Since, all the other ISPs have already bought port of capacity 1, they are not going to change their purchased port capacities as well (they will face the same

cost as Eq. 5), resulting in equilibrium. Hence, $X = \{1, 1, \dots, 1\}$ is an equilibrium solution, which results in a total cost of $C = N(2N + 2 + \epsilon)$.

On the other hand, if all of the ISPs were to buy a port capacity of 2, then the total cost would have been $N(2(2 + \epsilon) + 2(2N - 2N)) = N(4 + 2\epsilon)$, which is also the OPT solution. Then from the definition of *PoA* we have, $PoA \geq \frac{N(2N+2+\epsilon)}{N(4+2\epsilon)} = \frac{2N+2+\epsilon}{4+2\epsilon}$, which for large value of N approaches $N/2$.

Bounding *PoA* with Smoothness: For the scenario when C_S is dependent on X_i values, we use the smoothness property [14] to bound the *PoA*. According to the smoothness analysis if we can show that,

$$\sum_i [\lambda \cdot C_i(X^*) + \mu \cdot C_i(X) - C_i(X_i^*, X_{-i})] \geq 0, \tag{6}$$

then $PoA \leq \frac{\lambda}{1-\mu}$ holds for all X when X^* is the OPT.

Theorem 2. *The port capacity leasing game has a $PoA \leq 2$ if $y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i) \geq (y_i + y^i) C_S(X, y)$ holds for any value of $X, \forall i$.*

Proof. We prove the smoothness property (Eq. 6) for $\lambda = 1$ and $\mu = \frac{1}{2}$. Hence, the left hand side of Eq. 6 becomes,

$$\begin{aligned} & \sum_i \left[C_i(X^*) + \frac{1}{2} \cdot C_i(X) - C_i(X_i^*, X_{-i}) \right] \\ &= \sum_i y_i \left[C_P(X_i^*, y_i) + C_S(X^*, y) + \frac{1}{2} C_S(X, y) - C_S(X_i^*, X_{-i}, y) \right] \\ &+ \sum_i y^i \left[C_P(X_i^*, y^i) + C_S(X^*, y) + \frac{1}{2} C_S(X, y) - C_S(X_i^*, X_{-i}, y) \right] \\ &+ \frac{1}{2} \sum_i C_L(X_i). \end{aligned} \tag{7}$$

Now, when the assumption of Theorem 2 holds, we have $y_i C_P(X_i^*, y_i) + y^i C_P(X_i^*, y^i) \geq (y_i + y^i) C_S(X_i^*, X_{-i}, y) \forall i$. Then from Eq. 7 we always get a non-negative value and from smoothness argument we get, $PoA \leq \frac{1}{1-\frac{1}{2}} = 2$. \square

Corollary 1. *If both C_P and C_S represent M/M/1 delay functions, and the switch has a capacity of $\sum_i X_i$, then $PoA \leq 2$.*

Proof. This is actually a special case of Theorem 2. Since C_P represent M/M/1 delay function, $C_P(X_i, y_i) = \frac{1}{X_i - y_i}$, and $C_P(X_i, y^i) = \frac{1}{X_i - y^i}, \forall i$. Also, since the switch capacity is $\sum_i X_i$, $C_S(X, y) = \frac{1}{\sum_i [X_i - \frac{1}{2}(y_i + y^i)]}$. The multiplier of $\frac{1}{2}$ in front of $y_i + y^i$ is to avoid double counting, since y_{ij} is included in both y_i and

y^j . To prove the condition of Theorem 2 we observe that,

$$\begin{aligned}
 & y_i [C_P(X_i^*, y_i) - C_S(X_i^*, X_{-i}, y)] + y^i [C_P(X_i^*, y^i) - C_S(X_i^*, X_{-i}, y)] \\
 &= \left[\left(\frac{y_i}{X_i^* - y_i} + \frac{y^i}{X_i^* - y^i} \right) - \frac{y_i + y^i}{\sum_i (X_i^* + X_j + X_k + \dots) - \frac{1}{2}(y_i + y^i)} \right] \\
 &\geq \left[\left(\frac{y_i}{X_i^* - y_i} + \frac{y^i}{X_i^* - y^i} \right) - \frac{y_i + y^i}{\left(X_i^* - \frac{1}{2}(y_i + y^i) \right)} \right] \geq 0; \tag{8}
 \end{aligned}$$

where the first inequality is true because $X_i \geq \max(y_i, y^i), \forall i$; and $\frac{1}{a+b} < \frac{1}{a}$, for any $a, b > 0$. Then the last inequality can be proved by simple algebraic manipulation. Hence, the condition of Theorem 2 is met and we get a $PoA \leq 2$. \square

Finally, we argue why the assumption $y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i) \geq (y_i + y^i) C_S(X, y)$ is expected to hold even for fairly general C_P and C_S functions (beyond M/M/1 delay functions), as long as the switch is not the bottleneck in the traffic path through the IXP. Let $\bar{X} = \sum_i X_i$ denote the aggregate port capacity, and $\bar{y} = \frac{1}{2} \sum_i (y_i + y^i)$ denote the aggregate traffic through the IXP. Let us model the switch as a server with aggregate capacity \bar{X} and aggregate traffic \bar{y} , and with slight abuse in notation we denote the switch congestion (delay) cost as $C_S(\bar{X}, \bar{y})$. For any i , we have $\bar{X} - \bar{y} = X_i + \sum_{j \neq i} X_j - y_i - \sum_{j \neq i} y_j \geq X_i - y_i$, since $X_j \geq y_j \forall j$. As in typical queuing systems, the average delay is inversely proportional to the idle time $\bar{X} - \bar{y}$. Therefore, for any i , we can write, $C_S(\bar{X}, \bar{y}) \leq C_S(X_i, y_i)$. We further assume that the switch is not the bottleneck in the traffic path in terms of congestion cost, i.e., for any scalar values \tilde{X}, \tilde{y} , with $\tilde{X} \geq \tilde{y}$, we have $C_P(\tilde{X}, \tilde{y}) \geq C_S(\tilde{X}, \tilde{y})$. Taking $\tilde{X} = X_i, \tilde{y} = y_i$, we get $C_S(X_i, y_i) \leq C_P(X_i, y_i)$. Therefore, we have, $C_S(\bar{X}, \bar{y}) \leq C_S(X_i, y_i) \leq C_P(X_i, y_i)$. Similarly, we can argue, $C_S(\bar{X}, \bar{y}) \leq C_S(X_i, y^i) \leq C_P(X_i, y^i)$. Combining, we get $(y_i + y^i) C_S(\bar{X}, \bar{y}) \leq y_i C_P(X_i, y_i) + y^i C_P(X_i, y^i)$.

4 Simulations

4.1 Simulation Setup

Simulation of the port capacity leasing game requires: *a*) determining a viable traffic matrix (values of $y_{ij}, \forall (i, j)$ pair), and *b*) deciding the cost functions C_P, C_S, C_L such that they reflect real world behavior.

Traffic Matrix: To determine the traffic matrix, we individually calculated the y_{ij} values using a modified gravity model similar to the one used in [17]. We utilized PeeringDB to obtain information about IXP locations, the number of ISPs publicly peering at an IXP, and the port capacity purchased by the ISPs. On the other hand, utilized the CAIDA database [5] for router information (i.e., location and frequency at that location) of the ISPs. Using these information

we calculated the value of $y_i = \sum_{A,B} y_{i_A j_B}$, where $y_{i_A j_B}$ is the traffic that ISP i needs to send to ISP j from location A to location B . Using the gravitational law we have, $y_{i_A j_B} = k \frac{R_{i_A} \cdot R_{j_B}}{d_{AB}^2}$, where R_{i_A} is the number of Routers that ISP i has in area A , R_{j_B} is the number of Routers that ISP j has in area B , d_{AB} is the distance between area A and B , and K is a constant. The generated traffic matrix showed non-homogeneous traffic with exponentially decaying distribution, which supports the asymmetry usually observed in internet traffic [12].

4.2 Results and Observation

All of the simulation results discussed in this section are expressed in terms of PoS , calculated as the ratio of cost at any equilibrium to the OPT. Also, all the PoS results are for 4-ISP cases unless otherwise stated.

Shared Switch Delay: We model the switch cost C_S as the average packet delay through the switch and consider the following delay functions (C_S as a function of $\bar{X} = \sum_i X_i$ and $\bar{y} = \frac{1}{2} \sum_i (y_i + y^i)$): i) M/M/1 queuing delay function ($C_S = \frac{\beta}{\bar{X} - \bar{y}}$), ii) exponentially decaying delay function ($C_S = \beta \cdot \exp(-(\bar{X} - \bar{y}))$), and iii) linearly decaying delay function ($C_S = \beta \cdot (k - (\bar{X} - \bar{y}))$), where β is a proportionality factor and k is some constant. The effect on PoS for these three functions with the change of β is shown in Fig. 2. We observe that if we scale up C_S (increase β), then PoS generally becomes larger. The outcome is in agreement to the main idea of Theorem 2, where we argued that if C_S is small compared to the C_P , then PoS will be small.

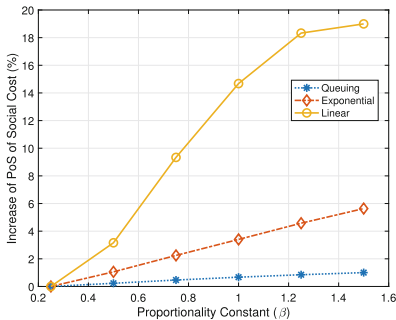


Fig. 2. Average PoS vs β for different switch congestion cost (delay) functions (C_S).

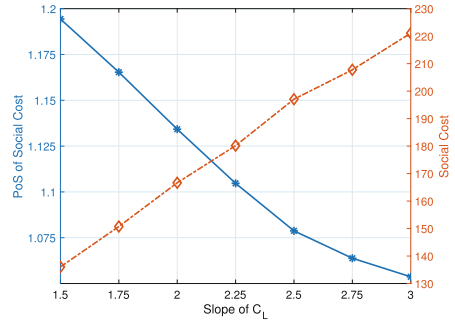


Fig. 3. Average PoS and SC vs per-unit leasing cost.

Port Leasing Cost: The effect of increasing the per unit leasing cost (C_L) is depicted in Fig. 3. We have assumed C_L to be linear in X_i and have increased the slope ($\frac{C_L(X_i)}{X_i}$) to characterize the increase of per unit port leasing cost. As we

can see, although increasing C_L decreases the PoS value of the system, it results in an increase in Social Cost (as defined in Eq. 2). Hence, although increasing the value of C_L will proportionally increase the revenue of IXP, it will also increase the social cost, eventually forcing the ISPs to look for alternate ways to exchange their traffic. The decrease of PoS with the increase of C_L can be explained in the following way: with the increase of C_L both the equilibrium and OPT solution results in very small values of X_i (because ISPs will try to buy as small port capacity as possible), thus making cost at equilibrium very close to cost at OPT.

Traffic Symmetry: The effect of traffic symmetry on PoS is measured by a factor $\alpha \in [0, 1]$ (coefficient of symmetry). The traffic values are generated using a uniform distribution with bounds $[\alpha \cdot y_{max}, y_{max}]$, where $y_{max} = \max_{i,j}(y_{ij}), \forall(i, j)$. Hence, the lower the value of α the higher the asymmetry. As shown by the simulation results in Fig. 4, the more symmetric the traffic ($\alpha = 1$ is homogeneous traffic) the higher the PoS . Also, this effect is more prominent when C_S is linear. Since internet traffic is usually not homogeneous [12], we can expect smaller PoS values in real world scenario.

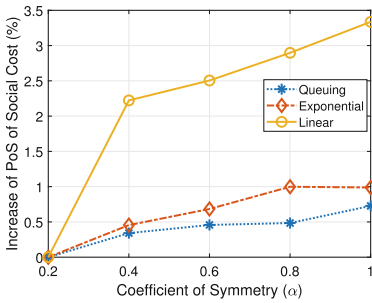


Fig. 4. Average PoS vs α for different switch congestion cost (delay) functions (C_S).

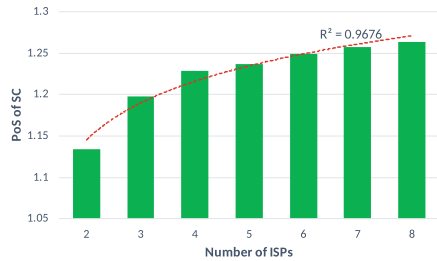


Fig. 5. Average PoS vs number of ISPs.

Number of ISPs: The change of PoS with the increase of ISPs doing public peering at some IXP is shown in Fig. 5. Since the computing time for finding the OPT solution is in the order of $O(k^N)$, where N is the number of ISPs, we limited our simulation to find the PoS values only for $N \leq 8$. From the results for $N = 2$ to 8, we observe that PoS increases with increasing N . Also, the variation of PoS against N seems to follow a logarithmic trend.

5 Conclusion

Port capacity leasing by ISPs at an IXP is modeled as a non-cooperative game, where all ISPs try to unilaterally minimize their individual costs. Theoretical

analysis of this game indicates that while there can be equilibria with very bad PoA , under reasonable practical assumptions PoA is bounded by 2. Simulation results also corroborate the findings of the theoretical results, and provide insight about the effect of congestion and port leasing cost functions on the PoS values. Traffic models generated using IXP data indicates high asymmetry of Internet traffic, for which the PoS values are observed to be very close to 1.

Acknowledgements. The authors should like to thank the National Science Foundation for supporting this work through award CNS-1816396.

References

1. Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S., Willinger, W.: Anatomy of a large European IXP. In: Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, pp. 163–174 (2012)
2. Alam, M.I.I., Anshelevich, E., Kar, K., Yuksel, M.: Proportional pricing for efficient traffic equilibrium at internet exchange points. In: 2021 33rd International Teletraffic Congress (ITC 33). IEEE (2021)
3. Anshelevich, E., Bhardwaj, O., Kar, K.: Strategic network formation through an intermediary. *Theory Comput. Syst.* **63**(6), 1314–1335 (2019). <https://doi.org/10.1007/s00224-018-09906-8>
4. Böttger, T., et al.: The elusive internet flattening: 10 years of IXP growth. arXiv e-prints (2018)
5. CAIDA: Macroscopic Internet Topology Data Kit (ITDK) (2020). <https://www.caida.org/catalog/datasets/internet-topology-data-kit/>
6. Calvó-Armengol, A., İlkılıç, R.: Pairwise-stability and Nash equilibria in network formation. *Internat. J. Game Theory* **38**(1), 51–79 (2009)
7. Cardona Restrepo, J.C., Stanojevic, R.: IXP traffic: a macroscopic view. In: Proceedings of the 7th Latin American Networking Conference, pp. 1–8 (2012)
8. Chiesa, M., Demmler, D., Canini, M., Schapira, M., Schneider, T.: SIXPACK: securing internet EXchange points against curious Onlookers. In: Proceedings of the 13th International Conference on Emerging Networking EXperiments and Technologies (CoNEXT 2017), pp. 120–133. ACM, New York, NY, USA (2017)
9. EuroIX: European IXP Reports (2021). <https://www.euro-ix.net/en/services/euro-ix-reports/>. Accessed 26 September 2021
10. Fabrikant, A., Luthra, A., Maneva, E., Papadimitriou, C.H., Shenker, S.: On a network creation game. In: Proceedings of the Twenty-Second Annual Symposium on Principles of Distributed Computing, pp. 347–351 (2003)
11. Jackson, M.O.: A survey of network formation models: stability and efficiency. *Group Formation Econ. Networks, Clubs, Coalitions* **664**, 11–49 (2005)
12. Labovitz, C., Iekel-Johnson, S., McPherson, D., Oberheide, J., Jahanian, F.: Internet inter-domain traffic. *ACM SIGCOMM Comput. Commun. Rev.* **40**(4), 75–86 (2010)
13. Norton, W.: *The Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press, London (2012)
14. Roughgarden, T.: Intrinsic robustness of the price of anarchy. In: Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing, pp. 513–522 (2009)

15. Snijders, J., Abdel-Hafez, S., Strong, M., Alom, C., Stucchi, M.: IXP Megabit/sec cost & comparison. <http://peering.exposed/>. Accessed 26 September 2021
16. Tardos, E., Wexler, T.: Network formation games and the potential function method. *Algorithmic Game Theory*, 487–516 (2007)
17. Zhang, Y., Roughan, M., Lund, C., Donoho, D.L.: Estimating point-to-point and point-to-multipoint traffic matrices: an information-theoretic approach. *IEEE/ACM Trans. Networking* **13**(5), 947–960 (2005)



A Phase Transition in Large Network Games

Abhishek Shende¹(✉), Deepanshu Vasal², and Sriram Vishwanath¹

¹ University of Texas, Austin, USA
{ashende, sriram}@utexas.edu

² Northwestern University, Evanston, USA
dvasal@umich.edu

Abstract. In this paper, we use a model of large random network game where the agents play selfishly and are affected by their neighbors, to explore the conditions under which the Nash equilibrium (NE) of the game is affected by a perturbation in the network. We use a phase transition phenomenon observed in finite rank deformations of large random matrices, to study how the NE changes on crossing critical threshold points. Our main contribution is as follows: when the perturbation strength is greater than a critical point, it impacts the NE of the game, whereas when this perturbation is below this critical point, the NE remains independent of the perturbation parameter. This demonstrates a phase transition in NE which alludes that perturbations can affect the behavior of the society only if their strength is above a critical threshold. We provide numerical examples for this result and present scenarios under which this phenomenon could potentially occur in real world applications.

Keywords: Game theory · Network games · Phase transition

1 Introduction

The ever increasing interactions between people in the world has motivated the analysis of network data in a range of disciplines and applications, appearing in such diverse areas as commerce, sociology, epidemiology, computer science, and national security. Network data is characterized by the edges between nodes and edge weights, through an adjacency matrix, wherein the action of individual player is affected by the actions of its neighbors. These actions not only affect the individuals but also the overall society. These outcomes can be modeled and studied through the framework of network games, focusing on understanding the effect of properties of network on Nash equilibria (NE). The authors in [29] and [16] study the properties of Nash equilibrium in different types of network games and how these are impacted by network structure.

A majority of real-world networks are large networks, with numerous players and interactions. Many random matrix models have been created that incorporate features of real-world systems, such as, network composition, flexibility,

and recurrent motifs [15, 28]. These network models appear in the study of complex systems, such as social networks, economic markets, signal processing and natural ecosystems. Characterizing the conditions on such networks for Nash equilibrium, even approximately, helps in understanding the results when players play selfishly. For large network games, we assume that the adjacency matrix is a random matrix, generated through a given distribution.

In statistical physics, the Ising model on large random graphs helps to study the phase transition phenomenon observed in the real world. It is shown in [20, 30], that the critical point behavior occurs for ferromagnetic interactions, where the magnetization vanishes as the continuously increasing temperature crosses a certain threshold. The phase transition between water and steam is an example of a phase transition occurring at a critical, or Curie temperature. This transition can be modeled with the Ising model, where the nodes are points in space, and the presence or absence of a molecule is generated randomly, like spins. Thus the magnetism corresponds to the density of the H_2O . At a hundred degrees Celsius, the density changes substantially; a phase transition. The author Malcolm Gladwell in [13] argues that there exist similar ‘tipping points’ in the society that are critical moments when a minor change makes all the difference. Tipping points are derived from ideas in epidemiology, the study of the spread of viruses and other diseases. The example of a simplified flu epidemic, as provided by Gladwell, describes how the start of Christmas season is a tipping point which leads to exponential increase in transmission rate due to such a simple cause as crowded Christmas shopping and cold weather. The theory of epidemic transmission is applicable to many ideas, products, messages, and behaviors we find in society can be characterized by their rapid, exponential spread through our population. The resurgence of brands like Hush Puppies and viral growth of new ones like Airwalk is contributed to early adoption and focus of marketing on influencers in the social network which the book calls as connectors, mavens, and salesmen. Through our paper, we try to use a game theoretic notion to mathematically study how tiny shift across the threshold can create huge effects by propagating the cause throughout the network.

In this paper, we explore a phase transition phenomenon seen in large random matrices in the context of a network game, and observe how the Nash Equilibrium of a game changes on varying a certain parameter. The authors in [7] have studied the extreme eigenvalues and eigenvectors of finite, low rank perturbations of random matrices. A phase transition phenomenon occurs whereby the large matrix limit of the extreme eigenvalues of the perturbed matrix differs from that of the original matrix if and only if the eigenvalues of the perturbing matrix are above a certain critical threshold. Our results focus on the ‘spiked’ random matrix models with Wigner and Wishart random ensemble for additive and multiplicative perturbations respectively. The key to applying the Baik, Ben Arous and Peche - BBP phase transition results from [3] lies in being able to compute the Cauchy or T transforms of the probability measure and their associated functional inverses. We use results on Wigner studied in [6, 11, 12, 14] and Wishart studied in [3, 4, 17, 25] random ensemble where the transforms and their

inverses can be expressed in closed form. Through previous work in [8, 26, 29] on how network structure affects the properties of Nash Equilibrium, we know that largest eigenvalue/eigenvector of the adjacency matrix plays a major role in determining the conditions of various properties of Nash Equilibrium of Linear Quadratic (LQ) network games. In this paper, we analyse perturbations to the adjacency matrix of large network games, focusing on how the changes to structure of network leads to a phase transition phenomenon of the NE of the perturbed network, due to its dependence on the extreme eigenvalue/eigenvector. To the best of our knowledge this is the first paper that makes a connection between random matrix theory and network games to show a phase transition in large social networks.

The application of such property can be experienced in different domains. The social media enables the communication among people, and promotes several group activities in the society. Based on data from platforms like Facebook, Twitter, etc., some users like celebrities, athletes, or politicians have significantly more followers than the rest. In a network graph terms, these users are the nodes that have much more influence than the rest of nodes. These social media influencers have a loyal follower base, achieving a high level of engagement on their content, such as images, trends, videos, etc., heightening their power of persuasion. Using our analysis, we can see that by introducing a perturbation through a ‘teleportation’ term into the adjacency matrix of a random network over a critical threshold, the beliefs of the participants in the network can depend on properties of the perturbation strength. This change in beliefs can lead to adoption of new technology or products through changing the influence of neighbors.

Some examples of network games with linear-quadratic model to observe a phase transition are public goods game [1, 27], influence of peers in education [5, 9] and to model criminal social interactions [10]. Crime and delinquency are related to connections in social networks, where delinquents often have friends who have committed offenses, and social ties are a means of influence to commit crimes. The ‘tipping point’ concept of using perturbation to affect the NE suggests that the properties of friendship networks should be taken into account to better understand peer influence on delinquent behavior and to craft delinquency-reducing policies.

Another application of the linear quadratic model is to model collaboration between firms as presented in [19]. Collaboration takes a variety of forms which includes creation and sharing of knowledge about markets and technologies, setting market standards and sharing facilities. The effects of peers has been evident in case studies on the adoption of high yielding hybrid crops by the farmers during the Dust Bowl in USA in late-1920s and 1930s [22]. The farmers were reluctant and slow to adopt hybrid crops, largely contributing to expensive switch and very few neighbors using hybrid crops. It was the focus by the government on young farmers over older farmers in combination with outreach by the USDA (headed by hybrid corn pioneer Henry A. Wallace), which can be interpreted as external perturbation on the existing network, that convinced Midwestern farmers to adopt the new seed [23, 24].

2 Nash Equilibrium in LQ Game

A network game \mathcal{G} with set of N players, is played over a weighted directed network whose structure is captured by an $n \times n$ adjacency matrix G . The (i, j) th entry of G , denoted by g_{ij} , represents the strength and type of influence of player j 's strategy on the utility function of player i . The positive (negative) g_{ij} represent strategic complements (substitutes) where an increase in neighbor j 's actions leads to a corresponding increase (decrease) in player i 's action. The action of a player i is given by $x_i \in \mathcal{X}_i \subseteq \mathbb{R}_{\geq 0}$, and so $x = (x_1, x_2 \dots x_n) \in \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i$. Each player $i \in \mathbb{N}[1, N]$ chooses their action $x_i \in \mathbb{R}_{\geq 0}$ to maximize a utility function:

$$J_i(x_i, z_i(x)),$$

which in turn depends on their own action x_i and on the aggregate neighbors' strategies $z_i(x)$, defined by the weighted linear combination

$$z_i(x) = \sum_{j=1, j \neq i}^N g_{ij}x_j.$$

The best response for player i , i.e., the action that maximizes the utility function is defined as

$$B_i(z_i(x)) := \arg \max_{x_i} J_i(x_i, z_i(x)).$$

The set of actions within which no player has an incentive for unilateral deviations (i.e., each player is playing a best response to other player's actions) is a Nash equilibrium. Mathematically, a vector $x^* = (x_1^*, \dots x_n^*)$, is a Nash equilibrium (NE) if, for all players $i \in \mathbb{N}[1, N]$, $x_i^* \in B_i(z_i(x))$.

A linear quadratic (LQ) network game is one where each agent chooses a scalar strategy $x_i \geq 0$ in order to maximize the linear quadratic utility function:

$$J_i(x_i, z_i(x)) = [z_i(x) + a_i]x_i - \frac{1}{2}q_i(x_i)^2 \tag{1}$$

with $a_i, q_i \in \mathbb{R}$.

The Nash equilibrium (NE) for an LQ game can be derived from the first-order necessary condition to maximize the utility function for each player i given by:

$$\frac{\partial J_i(x_i, z_i(x))}{\partial x_i} = \sum_{j=1, j \neq i}^N g_{ij}x_j + a_i - q_i x_i = 0. \tag{2}$$

This results in

$$q_i x_i^* = a_i + \sum_{j=1, j \neq i}^N g_{ij}x_j^*. \tag{3}$$

2.1 Assumption

To prove our main results, we require assumptions on the linear quadratic utility function in (1). In order to proceed, we impose the assumption that the vector $a = (a_1, \dots, a_n) = 0$ and $q_i = \lambda_{max} - g_{ii}$, where λ_{max} is the maximum eigenvalue of the adjacency matrix G .

Using these assumptions, (3) in matrix form, is

$$(\lambda_{max}I - G)x^* = 0, \tag{4}$$

where I is $N \times N$ identity matrix

$$\lambda_{max}x^* = Gx^* \tag{5}$$

Thus, we see that the NE x^* is the eigenvector of G corresponding to the maximum eigenvalue λ_{max} .

For an eigenvector to be a valid NE strategy, it has to satisfy the condition $x_i \in \mathbb{R}_{\geq 0}$ where x_i is each element of the eigenvector. Throughout this paper, we only consider perturbed adjacency matrices that are symmetric and element wise positive, which guarantees that the eigenvector corresponding to maximum eigenvalue has positive entries, through Perron-Frobenius theorem.

3 Additive Perturbation in LQ Games

Our model for the adjacency matrix of a large network is in the category of the random matrices with fixed-rank deformation, which includes the signal-plus-noise model as typical example. A vast amount of work has been devoted to understanding the limiting behavior of the extreme eigenvalues and the associated eigenvectors of the deformed models. Since the seminal work of authors in [3], there has significant study to understand that the extreme eigenvalues undergo a so-called BBP phase transition along with the change of the strength of the deformation. There exists a critical threshold such that the extreme eigenvalue of the perturbed matrix will be within the end points of the spectral distribution if the strength of the deformation is less than or equal to the threshold, and will otherwise be outside of support of the limiting spectral distribution.

Assume X be an $n \times n$ symmetric Gaussian Wigner matrix, with independent, zero mean normal distribution with variance $\sigma^2/2n$ on off diagonal and σ^2/n on diagonal. It has been shown that the spectral measure of X converges almost surely to the well known semi-circle distribution with density

$$d\mu_x(x) = \frac{\sqrt{4\sigma^2 - x^2}}{2\sigma^2\pi} dx \text{ for } x \in [-2\sigma, 2\sigma]. \tag{6}$$

Henceforth, $\xrightarrow{\text{a.s.}}$ denotes almost sure convergence. The result from [2] shows that the extreme eigenvalues converge almost surely to the endpoints of the support i.e. $\lambda_{max}(X) \xrightarrow{\text{a.s.}} 2\sigma$.

Let perturbation matrix P , be a $n \times n$ symmetric matrix with rank r . Its eigenvalues are $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$. By phase transition theorem, as $n \rightarrow \infty$ we have, for

$$\tilde{X} = X + P, \tag{7}$$

$$\lambda_{max}(\tilde{X}) \xrightarrow{\text{a.s.}} \begin{cases} \theta_1 + \frac{\sigma^2}{\theta_1} & \text{if } \theta_1 > \sigma \\ 2\sigma & \text{otherwise} \end{cases} \tag{8}$$

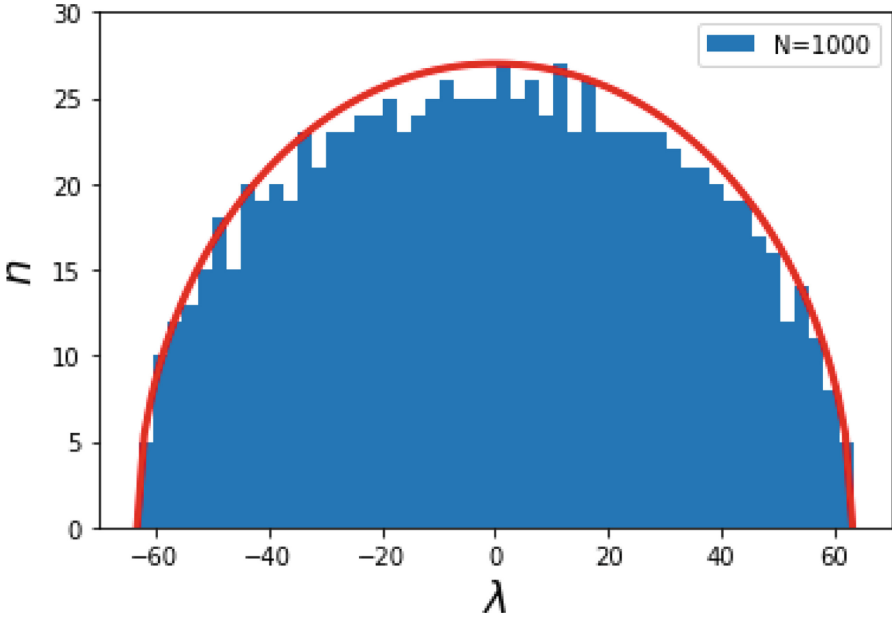


Fig. 1. Histogram of eigenvalues of Wigner matrix. The red curve represents the density of the semi-circle distribution (Color figure online)

Figure 1 shows the blue histogram of the eigenvalues of a Wigner matrix with $n = 1000$, where the red curve follows the semi-circle law from (6). The extreme eigenvalues are within the boundaries of the semi-circle. The blue histogram of the Fig. 2 is that of the eigenvalues of the same Wigner matrix but perturbed this time by a symmetric rank 1 matrix as in (7). The outlier represented here by blue dot outside red curve is exhibited by the extreme eigenvalue of deformed matrix out of the bulk spectrum. This shows the phase transition phenomenon from (8) when the $\theta > \sigma$.

In the setting where $r = 1$ and $P = \theta uu^T$, let \tilde{u} be a unit-norm eigenvector of \tilde{X} associated with its largest eigenvalue. The parameter θ represents the signal to

noise ratio. By eigenvector phase transition theorem, for norm of the eigenvector projection we have,

$$|\langle \tilde{u}, u \rangle|^2 \xrightarrow{\text{a.s.}} \begin{cases} 1 - \frac{\sigma^2}{\theta^2} & \text{if } \theta \geq \sigma \\ 0 & \text{if } \theta < \sigma \end{cases} \quad (9)$$

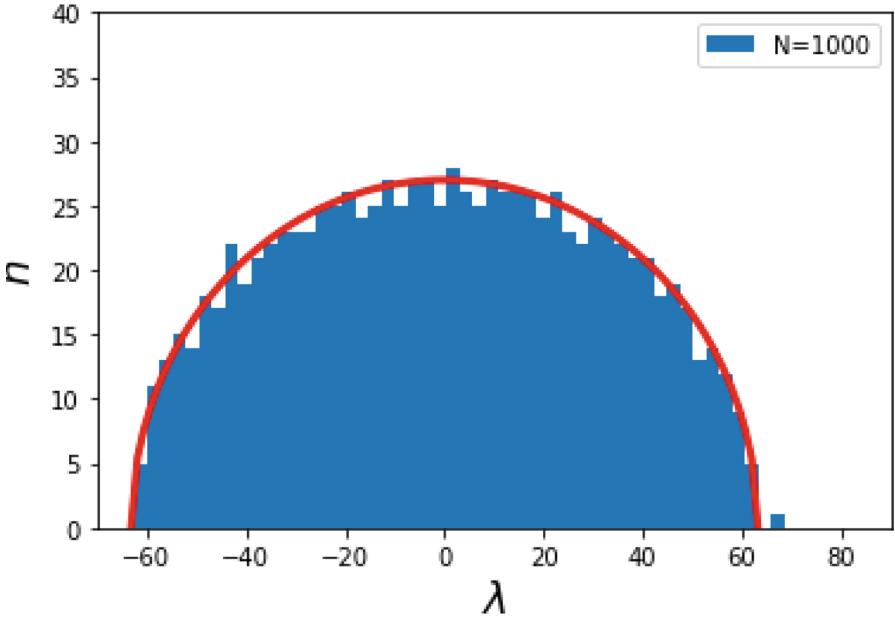


Fig. 2. Histogram of eigenvalues of rank 1 perturbation of Wigner matrix. The red curve represents the density of the semi-circle distribution (Color figure online)

3.1 Transitions in LQ Game

In the context of LQ game, the matrix X is the $n \times n$ symmetric Gaussian Wigner matrix noise that models the interactions between different participants of a large network. Let $P = \theta uu^T$ so that θ and u are largest eigenvalue and eigenvector respectively, be the adjacency matrix of the perturbations, leading to deviations in the impact of interactions between different players.

Considering an additive perturbation as in (7), the adjacency matrix of the new network is \tilde{X} . From (5), since the NE of the game described in (1) is the eigenvector of the adjacency matrix corresponding to the maximum eigenvalue, the NE for game with \tilde{X} is $\tilde{x}^* = \lambda_{max}(\tilde{X}) = \tilde{u}$.

From the result of (9), for the LQ game, the Nash Equilibrium of the perturbed game is dependent on the parameter θ of the perturbation. A transition

phenomenon is observed for values of θ greater than the threshold value of σ when the NE given by \tilde{u} is dependent on the network parameters i.e. θ and σ . Below this threshold value, the NE does not depend on the strength of perturbation and cannot be affected as desired by modifying the strength of network connections.

3.2 Numerical Examples

To demonstrate the transition in NE, we use a numerical example with $n = 2000$ Wigner matrix for additive perturbation as shown in (7). The value of θ is varied to see the effect of the threshold value of σ on the Nash equilibrium. Since the NE of our LQ game is the leading eigenvector, we use the product in (9) as the theoretical value and compare it with the computational result. The vertical dashed line in Fig. 3 and 4 is the critical transition point, across which the strength of the perturbation starts affecting the Nash equilibrium of the game.

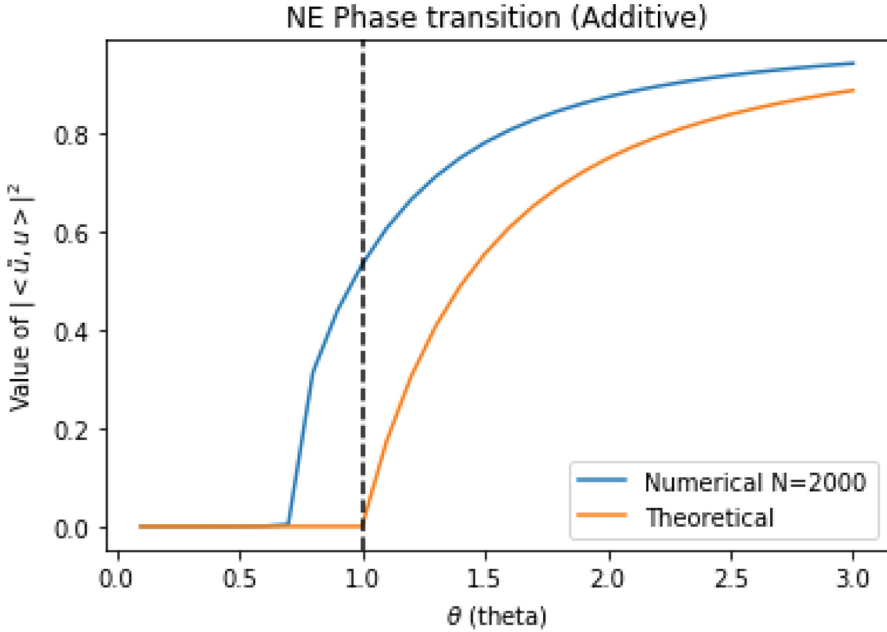


Fig. 3. Comparison of transition in NE for additive perturbation of LQ game with $\sigma = 1$.

In Fig. 3, the critical threshold value of $\sigma = 1$. The value of $|\langle \tilde{u}, u \rangle|^2$ from (9) is denoted by the ‘theoretical’ curve. The ‘numerical’ curve is derived by computationally calculating the NE, i.e. the leading eigenvector for the perturbed adjacency matrix.

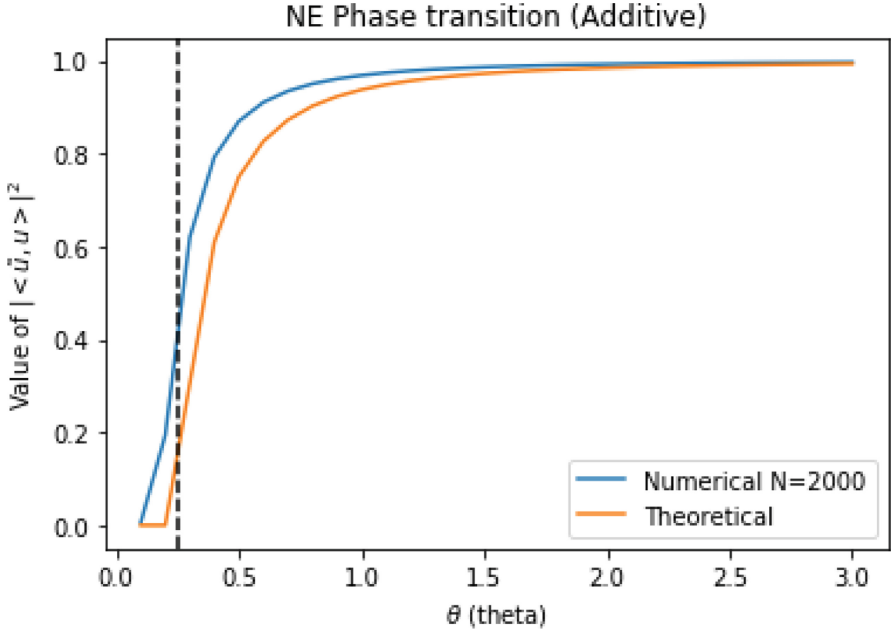


Fig. 4. Comparison of transition in NE for additive perturbation of LQ game with $\sigma = 0.25$.

In Fig. 4, the critical threshold value of $\sigma = 0.25$. The numerical and the theoretical curves follow closely. The minor variation in the values of the two curves occurs due to the fact that (9) is for range of $n \rightarrow \infty$.

4 Multiplicative Perturbation in LQ Game

Now we explore another strategy to modify a network, through a multiplicative deformation. Assume M_n be an $n \times m$ matrix with independent, zero mean, normally distributed entries with variance 1. Then, $X_n = M_n M_n^* / m$ is known as Wishart matrix. It has been shown that as $n, m \rightarrow \infty$ with $n/m \rightarrow c > 0$, the spectral measure of X_n converges almost surely to the well-known Marchenko-Pastur distribution [21] with density

$$d\mu_x(x) = \frac{1}{2\pi cx} \sqrt{(b-x)(x-a)} 1_{[a,b]}(x) dx + \max(0, 1 - \frac{1}{c}) \delta_0 \quad (10)$$

where $a = (1 - \sqrt{c})^2$ and $b = (1 + \sqrt{c})^2$. It is known that the extreme eigenvalues converge almost surely to the endpoints of this support.

Let perturbation matrix P , be a $n \times n$ symmetric matrix with rank r . It's eigenvalues are $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$. By phase transition theorem, as $n \rightarrow \infty$ we have, for

$$\tilde{X} = X(I + P) \tag{11}$$

$$\lambda_{max}(\tilde{X}) \xrightarrow{\text{a.s.}} \begin{cases} (\theta_1 + 1)(1 + \frac{c}{\theta_1}) & \text{if } \theta_1 \geq \sqrt{c} \\ (1 + \sqrt{c})^2 & \text{otherwise} \end{cases} \tag{12}$$

In the setting where $P = \theta uu^T$, let \tilde{u} be a unit-norm eigenvector of \tilde{X} from (11), associated with its largest eigenvalue. For Wishart matrix X as described above, the eigenvector phase transition occurs as,

$$|\langle \tilde{u}, u \rangle|^2 \xrightarrow{\text{a.s.}} \begin{cases} \frac{\theta^2 - c}{\theta[c(\theta + 2) + \theta]} & \text{if } \theta \geq \sqrt{c} \\ 0 & \text{if } \theta < \sqrt{c} \end{cases} \tag{13}$$

4.1 Transitions in LQ Game

In an LQ game, from (1), the large random matrix X is considered to be $n \times n$ random Wishart matrix, with $P = \theta uu^T$ is the adjacency matrix of the perturbations so that θ and u are largest eigenvalue and eigenvector respectively. Considering a multiplicative perturbation as in (11), the adjacency matrix of the new network is \tilde{X} . As seen in (5), the NE of \tilde{X} is given by the eigenvector corresponding to maximum eigenvalue. Thus the NE $\tilde{x}^* = \lambda_{max}(\tilde{X}) = \tilde{u}$.

Using the result from (13), for the LQ game, the Nash Equilibrium of the perturbed game is dependent on the parameter θ of the perturbation. A transition phenomenon is observed for values of θ greater than the threshold value of \sqrt{c} when the NE given by \tilde{u} is dependent on the network parameters i.e. θ and \sqrt{c} . Below this threshold value, the NE is not affected by the strength of the deformation and cannot be controlled by changing the edge weights of a network through external factors.

4.2 Numerical Examples

To demonstrate the transition of NE, we use a numerical example with $n = 2000$ Wishart matrix for the multiplicative perturbation as shown in (11). The value of \sqrt{c} , dependent on m as $n/m \rightarrow c$ is the threshold value given by dashed vertical line in Fig. 5 and 6. The value of θ is varied below and above \sqrt{c} to observe the effect on the Nash equilibrium. Since the NE of our LQ game is the leading eigenvector, we use the product in (13) as the theoretical value and compare it with the computational value.

In Fig. 5, the critical threshold value of $\sqrt{c} = 1$. The value of $|\langle \tilde{u}, u \rangle|^2$ from (13) is denoted by the ‘theoretical’ curve. The ‘numerical’ curve is derived by computationally calculating the NE, i.e. the leading eigenvector for the perturbed adjacency matrix. In Fig. 6, the critical threshold value of $\sqrt{c} = 1.414$. The numerical and the theoretical curve have differences which occur since in the numerical we have finite n, m whereas for the ‘theoretical’ we consider (13) which represents the limit $n, m \rightarrow \infty$. In both examples, we observe that the NE

depends on the strength of the deformation, when the strength is above the critical point there is phase transition for the NE.

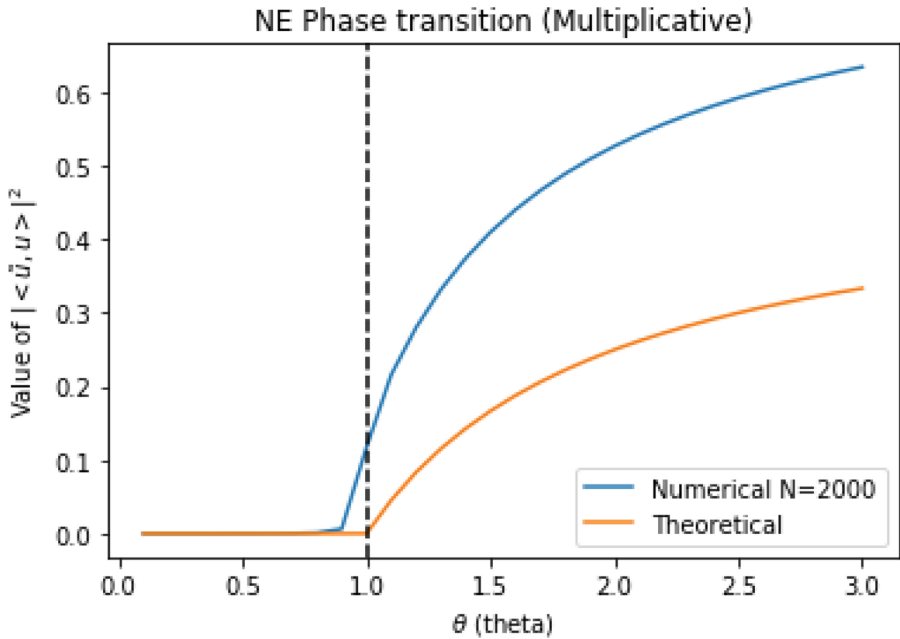


Fig. 5. Comparison of transition in NE for multiplicative perturbation of LQ game with $m = 2000$, thus $\sqrt{c} = 1$.

5 Interpretation

In the previous sections, we see how a large network can be influenced through external perturbations. The Nash Equilibrium, for the LQ game in (1), can be affected through additive and multiplicative perturbation. If the perturbation strength θ is strong, the primary eigenvalue goes beyond the random spectrum and the primary eigenvector, which is the NE of our LQ game is correlated with θ (in a cone around the perturbation’s eigenvector direction whose deviation angle goes to 0 as $\frac{\theta}{\sigma} \rightarrow \infty$). If θ is sufficiently low, the primary eigenvalue is buried in the random spectrum, and the NE is random, with no correlation to perturbation’s eigenvector.

The finite rank perturbation of an adjacency matrix of large network can be interpreted in multiple ways. The additive perturbation can be designed such the specific nodes of the network has higher influence through their edges on the neighbors and consequently the entire network. The multiplicative perturbation, though harder to visualize, also affects the edge weights to dominate the influence

from certain nodes and affect the NE properties. The NE, through eigenvalue phase transition is shown for rank 1 perturbation matrices. The rank 1 adjacency matrix can be designed in such a way that the influencers of a real world network have a major role to cause an aggressive spread over the network. The NE, which is the optimal participation of each member, is influenced by this perturbation. The shift in primary eigenvectors also affects the eigenvector centrality measure of the large networks.

The deformation to the networks can be manipulated such that the influence of some of the important nodes of the network is powerful enough to propagate their beliefs quicker and effectively. Work in the area of identifying the subset of individuals within a network that can maximize spread of idea has studied like in [18]. Using our analysis, we could apply sufficient external modifications to a network targeting influential set of individuals, to trigger a large cascade of belief spread.

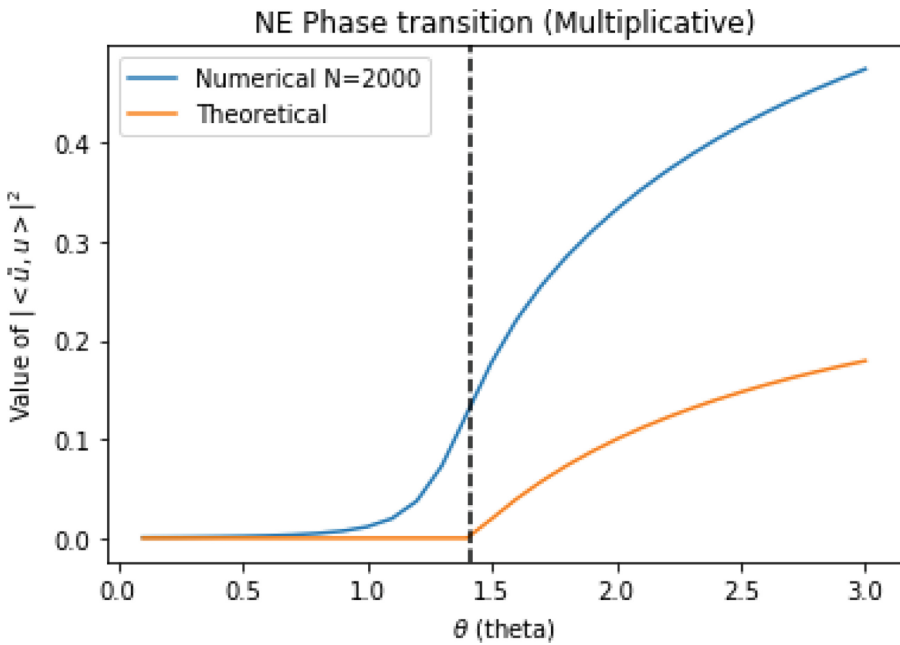


Fig. 6. Comparison of transition in NE for multiplicative perturbation of LQ game with $m = 1000$, thus $\sqrt{c} = 1.414$.

6 Conclusion

In this paper, we explore conditions under which the Nash equilibrium of a large network game undergoes a phase transition when the adjacency matrix of

the network is deformed through a finite matrix perturbation. Our contribution uses the well known eigenvector and eigenvalue phase transition phenomenon in conjunction with finite rank deformations of random matrices from field of random matrix theory to model such phenomena for linear quadratic games on a large network. For the LQ game, we make certain assumptions on the parameters of the utility function that enable us complete the analysis of transition of the NE. The large network is modelled as Wigner and Wishart random matrix for additive and multiplicative deformations respectively. The network property i.e. σ or c determines the critical point around which the strength of perturbation shifts the NE. We observe that, as θ crosses the critical point, the NE jumps out of the spectrum and is dependent on θ . For the values of θ below the critical point, the NE is unaffected by θ . There are multiple potential applications for this phenomenon, where selfish participants form a large network.

There exists several avenues for future work. In field of network games, multiple settings are described through variety of utility functions. There is possibility of studying such phase transition in other network games with non linear functions. The eigenvalue/eigenvector transition for random matrix deformation occurs for distributions other than Wigner/Wishart as well. It would be a new direction to investigate phase transition for complex systems that are modeled by other density functions. Our model uses rank 1 deformation but studying phase transition for higher rank deformations and what such perturbations would mean to the network is another possible avenue. The extreme eigenvalues and spectral properties of adjacency matrix play an important role on other properties of NE like conditions on uniqueness and stability. Exploring how sudden shift in extreme eigenvalues affects these properties can help in devising appropriate perturbations. To this end, our results provide a framework to tackle some of the above open problems and applications in future work.

References

1. Allouch, N.: On the private provision of public goods on networks. *J. Econ. Theory* **157**, 527–552 (2015). <https://doi.org/10.1016/j.jet.2015.01.007>
2. Anderson, G.W., Guionnet, A., Zeitouni, O.: *An Introduction to Random Matrices*. Cambridge University Press (2009). <https://doi.org/10.1017/cbo9780511801334>, <https://www.cambridge.org/core/books/an-introduction-to-random-matrices/8992DA8EB0386651E8DA8214A1FC7241>
3. Baik, J., Arous, G.B., Peche, S.: Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *Ann. Probab.* **33**(5), 1643–1697 (2004), <http://arxiv.org/abs/math/0403022>
4. Baik, J., Silverstein, J.W.: Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97**(6), 1382–1408 (2006). <https://doi.org/10.1016/j.jmva.2005.08.003>
5. Ballester, C., Calvó-Armengol, A., Zenou, Y.: Who's who in networks. Wanted: the key player. *Econometrica* **74**, 1403–1417 (2006). <https://doi.org/10.1111/j.1468-0262.2006.00709.x>
6. Bassler, K.E., Forrester, P.J., Frankel, N.E.: Eigenvalue separation in some random matrix models. *J. Math. Phys.* **50**(3), 033302 (2008). <https://doi.org/10.1063/1.3081391>, <http://arxiv.org/abs/0810.1554>, <https://dx.doi.org/10.1063/1.3081391>

7. Benaych-Georges, F., Nadakuditi, R.R.: The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices (2009). <http://arxiv.org/abs/0910.2120>
8. Bramoullé, Y., Kranton, R., Bramoullé, Y., Kranton, R.: Games played on networks. In: *The Oxford Handbook of the Economics of Networks* (2016). <https://doi.org/10.1093/oxfordhb/9780199948277.013.8>
9. Calvó-Armengol, A., Patacchini, E., Zenou, Y.: Peer effects and social networks in education. *Rev. Econ. Stud.* **76**, 1239–1267 (2009). <https://doi.org/10.1111/j.1467-937X.2009.00550.x>
10. Calvó-Armengol, A., Zenou, Y.: Social networks and crime decisions: the role of social structure in facilitating delinquent behavior (2004). <https://doi.org/10.1111/j.0020-6598.2004.00292.x>, <https://onlinelibrary.wiley.com/doi/full/10.1111/j.0020-6598.2004.00292.x>, <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0020-6598.2004.00292.x>, <https://onlinelibrary.wiley.com/doi/10.1111/j.0020-6598.2004.00292.x>
11. Capitaine, M., Donati-Martin, C., Féral, D.: The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37**(1), 1–47 (2009). <https://doi.org/10.1214/08-AOP394>, <https://projecteuclid.org/journals/annals-of-probability/volume-37/issue-1/The-largest-eigenvalues-of-finite-rank-deformation-of-large-Wigner/10.1214/08-AOP394.full>, <https://projecteuclid.org/journals/annals-of-probability/volume-37/issue-1/The-largest-eig>
12. Féral, D., Pécché, S.: The largest eigenvalue of rank one deformation of large Wigner matrices. *Commun. Math. Phys.* **272**(1), 185–228 (2007). <https://doi.org/10.1007/s00220-007-0209-3>, <https://link.springer.com/article/10.1007/s00220-007-0209-3>
13. Gladwell, M.: *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown and Co (2002)
14. Hoyle, D.C., Rattay, M.: Statistical mechanics of learning multiple orthogonal signals: Asymptotic theory and fluctuation effects. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **75**(1) (2007). <https://doi.org/10.1103/PhysRevE.75.016101>, <https://pubmed.ncbi.nlm.nih.gov/17358218/>
15. Jackson, M.O.: *Social and Economic Networks*. Princeton University Press, Princeton (2010)
16. Jackson, M.O., Zenou, Y.: Games on networks. In: *Handbook of Game Theory with Economic Applications* (2015). <https://doi.org/10.1016/B978-0-444-53766-9.00003-3>, <https://web.stanford.edu/jacksonm/GamesNetworks.pdf>
17. Karoui, N.E.: Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.* **35**(2), 663–714 (2005). <https://doi.org/10.1214/009117906000000917>, <http://arxiv.org/abs/math/0503109>, <https://dx.doi.org/10.1214/009117906000000917>
18. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146 (2003). <https://doi.org/10.1145/956750.956769>
19. Koenig, M.D., Liu, X., Zenou, Y.: R&D networks: theory, empirics and policy implications. discussion papers (2014). <https://ideas.repec.org/p/sip/dpaper/13-027.html>
20. Landau, D.P.: Finite-size behavior of the ising square lattice. *Phys. Rev. B* **13**(7), 2997–3011 (1976). <https://doi.org/10.1103/PhysRevB.13.2997>, <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.13.2997>

21. Marčenko, V.A., Pastur, L.A.: Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sbornik* **1**(4), 457–483 (1967). <https://doi.org/10.1070/sm1967v001n04abeh001994>
22. McLeman, R.A., Dupre, J., Berrang Ford, L., Ford, J., Gajewski, K., Marchildon, G.: What we learned from the Dust Bowl: lessons in science, policy, and adaptation. *Popul. Environ.* **35**(4), 417–440 (2014). <https://doi.org/10.1007/s11111-013-0190-z>, <http://link.springer.com/10.1007/s11111-013-0190-z>
23. Meyers, K., Rhode, P.W.: Exploring the causes driving hybrid corn adoption from 1933 to 1955. Technical report (2019)
24. Moscona, J., et al.: Environmental catastrophe and the direction of invention: evidence from the American dust bowl *. Technical report (2021). <http://economics.mit.edu/grad/moscona>
25. Nadler, B.: Finite sample approximation results for principal component analysis: a matrix perturbation approach. *Annals of Statistics* **36**(6), 2791–2817 (2008). <https://doi.org/10.1214/08-AOS618>, <https://projecteuclid.org/journals/annals-of-statistics/volume-36/issue-6/Finite-sample-approximation-results-for-principal-component-analysis-A/10.1214/08-AOS618.full>, <https://projecteuclid.org/journals/annals-of-statistics/volume-36/issue-6/Finite-sampl>
26. Naghizadeh, P., Liu, M.: On the uniqueness and stability of equilibria of network games. In: 55th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2017, vol. 2018-Janua, pp. 280–286 (2017). <https://doi.org/10.1109/ALLERTON.2017.8262749>
27. Naghizadeh, P., Liu, M.: Provision of public goods on networks: on existence, uniqueness, and centralities. *IEEE Trans. Netw. Sci. Eng.* (2018). <https://doi.org/10.1109/TNSE.2017.2755003>, <https://arxiv.org/pdf/1604.08910.pdf>
28. Newman, M.: *Networks*. Oxford University Press, Oxford (2018)
29. Parise, F., Ozdaglar, A.: A variational inequality framework for network games: existence, uniqueness, convergence and sensitivity analysis. *Games Econ. Behav.* **114**, 47–82 (2017), <http://arxiv.org/abs/1712.08277>
30. Stanley, H.E.: Scaling, universality, and renormalization: three pillars of modern critical phenomena (1999). <https://doi.org/10.1103/revmodphys.71.s358>, <https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.71.S358>



Dynamic Pricing for Tenants in an Automated Slicing Marketplace

Alessandro Lieto¹(✉), Iliaria Malanchini¹, Silvio Mandelli¹,
and Antonio Capone²

¹ Nokia Bell Labs, Stuttgart, Germany
{alessandro.lieto, ilaria.malanchini,
silvio.mandelli}@nokia-bell-labs.com

² Politecnico di Milano, Milan, Italy
antonio.capone@polimi.it

Abstract. The paradigm shift from a *one-size-fits-all* architecture to a *service-oriented* network infrastructure promised by network slicing will demand novel technical solutions, as well as new business models. In particular, the role separation between infrastructure providers, i.e. the ones owning the network, and slice tenants, i.e. the ones providing specialized services tailored to their vertical segments, may encourage the definition of a shared platform (or marketplace) where the former can monetize their network infrastructure by leasing network resources at a market price, and the latter can rent on-demand the network resources needed to offer their services at the desired quality. This also enables the flexibility for the slice tenants to optimize the management of their slices by adapting their resource demand to fluctuations of their traffic or variations of the price in the market. In this paper, we extend the market mechanism scheme developed in previous works by including intra-slice radio admission control policies in the utility definition of the tenants in the slicing market game. Moreover, we characterize the mathematical properties of the game with respect to slice configuration, i.e. how diverse strategical behavior of the tenants affects the market operation, in terms of slice resource allocation and performance. Our analysis offers insights to the slice tenants on how they could reconfigure their techno-economic performance indicators in response to the dynamics of network and of the market, namely how to adapt their long-term (and/or real-time) strategies to the fluctuations of the traffic to enhance network performance and increase profits.

Keywords: Network slicing · Game theory · Dynamic pricing · Slicing marketplace

1 Introduction

The key differentiator of the upcoming 5G and beyond systems from previous wireless technologies is the integration of vertical industries in the telcos ecosystem, which paves the way for new market opportunities and innovative business

models [2]. In particular, *network slicing* enables an Infrastructure Provider (INP) to support diverse services over a common network infrastructure by offering customized end-to-end (E2E) logical networks, i.e., slices, by sharing the same pool of network resources and functionalities. Once in operation, the slices may need to be dynamically scaled up/down to match any variation of service requirements or adapt their configuration to dynamic changes and/or undesired trend in their monitored Key Performance Indicators (KPIs). Therefore, openness of the network to third parties and flexibility in the management of the slices are key features to encourage vertical players to use existing network infrastructures rather than deploying their own private infrastructure. In [19], the authors define a framework based on a vertical-oriented network slicing design where slice tenants can entirely customize and upgrade their network slices, with zero-touch service and network management. Along the same lines, [7] addresses the key challenges of the life-cycle management of network slices, discussing the trade-off between the degree of control and customization of network slices (by the vertical tenant) and the operational complexity in the network management (by the INP), specifically for the management of the shared resources in Radio Access Network (RAN) [17], e.g., the spectrum. Due to the random nature of traffic, physical radio resource reservation might not suffice to provide neither resource efficiency nor expected guarantees in terms of Quality of Service (QoS) and Quality of Experience (QoE) [14, 15]. To maintain satisfactory user experience and high profits for slice tenants in a dynamic environment, a slice may need to be reconfigured according to the varying traffic demand and resource availability. However, existing works focus either on static resource allocation, leaving to admission control the decision of accepting a network slice to guarantee the Service Level Agreement (SLA) in varying traffic conditions [6, 16], or defines centralized optimization routines to dynamically allocate the resources to increase resource efficiency or maximize social welfare [9, 11, 18]. However, since these slices will be used by profit business entities, e.g. verticals, the allocation of physical resources to the network slices must consider their private revenue and business models, in addition to the provisioning of desired QoS.

In [12, 13], we introduce a Slicing Management Framework (SMF) that can be applied for the dynamic orchestration of network resources owned by an INP in a multi-tenant shared marketplace. In those works, the market game with its mathematical properties is described, together with an algorithmic implementation to guarantee the convergence to a Nash Equilibrium (NE). In this paper we aim at completing the picture by testing the applicability of the SMF in a realistic scenario and providing additional insights on the impact of each tenant's choice in the evolution of the market due to the configuration of their slices. We show that the system automatically scales and adapts the resource allocation to the dynamics of a real network, being able to react to both short-term effects, like mobility and handovers, and long-term changes in the overall offered traffic at different times of a day, whose variations are taken from traffic traces of an LTE network collected in a urban area. Given the fluctuations of slices' traffic, we extend the analysis of the slicing market game by introducing

intra-slice Radio Admission Control (RAC) policies in the strategy evaluation of the tenants, as a mean to control slice congestion that can affect both network performance and costs for the resources, similarly to what is proposed in [20]. Moreover, differently from [12,13], we also offer insights to the slice tenants for a better customization of their network slice techno-economical configuration. This would allow them to better meet their business intents and desired slice's SLAs.

The paper is structured as follows. In Sect. 2, we describe the system model of the slicing marketplace and introduce the payoff functions of the game when including RAC. In Sect. 3, we derive some relevant theoretical properties of the game and analyze the impact of the slice configurations on their strategic behavior in the market. Numerical evaluation and simulation setup are discussed in Sect. 4 and, finally, Sect. 5 concludes the paper with final remarks and future works.

2 Slicing Market Game

We consider a network slicing setup consisting of a set of network slices, S , each one owned by a single tenant or vertical enterprise¹, that engage in a shared marketplace for purchasing radio network resources, i.e., spectrum, to be assigned to their slices. We further assume that the marketplace is controlled and managed by a single INP, which applies dynamic pricing policies to monetize the utilization of the network resources, e.g., by adapting the price for a unit of resources according to the current demand. Let $x_s \in [0, 1]$ be the resource allocation request of a slice s , normalized according to the total available system bandwidth, and let $l = \sum_{s \in S} x_s$ be the current load in the network, namely the total resource demand from all the slices. Then, we define the pricing function as

$$P(l) = 1 + e^{\gamma(l-l_0)}, \quad (1)$$

where parameters γ and l_0 allow one to tune the shape of the pricing function. In particular, when the total load approaches the value l_0 , the price increases exponentially to discourage the tenants to purchase extra resources. Accordingly, we compute the costs of the tenants for purchasing x_s resources in the market as

$$C_s(x_s, l) = x_s \cdot P(l), \quad (2)$$

where the term x_s depends solely on the tenant decision, while $P(l)$ accounts for the aggregate decisions of all the tenants (in terms of load). As done in [12,13], we model the interactions of the tenants in the marketplace as a *slicing market game*. Then, let $\mathcal{G} = \langle S, (X_s)_{s \in S}, (u_s)_{s \in S} \rangle$ be the strategic form of the slicing market game with S as the set of players (the slice tenants), $X_s \in [0, 1]$ the strategy space (normalized amount of resources to buy) and u_s the payoff functions of the players. In previous works [12,13], we analyze the property of

¹ This assumption holds for simplicity. In general, a single tenant may control multiple network slices and still have different business models for each of them.

the game, by showing that the game \mathcal{G} belongs to the subclass of *aggregative games* [8], given the aggregative property of the cost function $C_s(\cdot)$. We further prove the existence of NE for a class of utility (payoff) functions and provide an algorithmic implementation that guarantees the convergence to at least one of them, also discussing the quality of the achieved equilibria. We assume the payoff function of the players to consist of two terms:

- a private term, $R_s(x_s)$, describing the business model of the tenants that computes the expected returns given the achieved slice performance, which depends solely on the individual action of the slice tenant, and
- a costs term, $C_s(x_s, l)$, that varies according to the law of demand and supply of the market.

Therefore, the payoff function of the tenant can be computed as

$$u_s(x_s, \mathbf{x}_{-s}) = R_s(x_s) - C_s(x_s, \mathbf{x}_{-s}), \quad (3)$$

where we replace the load l with \mathbf{x}_{-s} to highlight the dependence on the other players' strategies. In particular, the revenue function of the tenants is expressed as

$$R_s(x_s) = \sum_{k \in K_s} \chi_k \cdot A_k(x_k), \quad (4)$$

where K_s is the set of users within a slice s , χ_k is a per-user economic value that the tenant assigns to each user k , and $A_k(x_k)$ is an acceptance probability function that quantifies the level of satisfaction of a user for the experienced quality of service. Notice that the value x_k defines the per-user resource allocation, which depends on how the radio scheduler redistributes the x_s resources of a slice among its users². We model the experienced quality of service (in terms of achieved throughput) of a user through the acceptance probability function defined in [5]:

$$A_k(x_k; \mu, r_0) = 1 - q \left(\frac{r(x_k)}{r_0} \right)^\mu, \quad (5)$$

where $q \in (0, 1)$, μ defines the steepness of the curve, i.e., the sensitivity of the user to performance degradation, r_0 represents the maximal throughput and $r(x_k) = x_k \cdot \eta_k$ denotes the achievable throughput given the allocated resources, x_k , and the experienced spectral efficiency, η_k . As in [12], we assume that the marginal utility shows diminishing return when approaching some reference value (i.e., maximal throughput r_0), namely that the increase in quality of experience of a user vanishes when approaching the requested QoS.

2.1 Admission Control Policies

In this paper, we extend the analyses of the previous works by including customized RAC policies for each slice. Indeed, the admission of new users in the

² In the rest of the work, we assume that the resources are evenly split among the users, i.e. $x_k \simeq \frac{x_s}{|K_s|}$ - with $|\cdot|$ denoting the cardinality of a set - as achieved by the state of the art proportional fair schedulers [4, 10].

system for a slice may require a renegotiation of the current purchased network resources to avoid degradation of performance of already admitted users. However, depending on the current load of the network, the purchase of extra resources might not necessarily lead to an increase in performance (or profits, due to higher prices). Therefore, the admission of new users can impact the decisions of the tenants in the market. Hence, we replace the payoff functions presented in Eq. (3) with the following formulation

$$u_s(x_s, n_s, \mathbf{x}_{-s}) = R_s(x_s, n_s) - C_s(x_s, \mathbf{x}_{-s}), \quad (6)$$

where the RAC decision of a tenant translates into the number of users admitted in the system, n_s . Notice that at time of resource negotiation, we assume the tenants to have complete knowledge of their current offered traffic. This assumption holds for tractability reason. However, this approach can be applied and further optimized predicting (e.g. by means of time series) their daily traffic, taking into account that, as a consequence, their RAC policies might be affected by potential prediction errors [6]. Given the complexity of this prediction use case (and the potential impact on its decisions as well as on the decisions of other players in the market game), we focus on a simplified scenario where the amount of traffic is known by the tenants and no prediction errors must be taken into account.

The RAC policy considered hereafter assumes that users arrive in the system at random time and, therefore, the tenants cannot select the users to be accepted according to favorable channel conditions, neither discriminate users based on their location (e.g., reject an user at cell edge). Indeed, the tenant evaluates one user per time and accept a new user only if it generates an increase in its profit function, $u_s(\cdot)$. After a new user is rejected, all following users are also rejected. The optimization variables of the tenants become, therefore, the tuple (x_s, n_s) , where the RAC policy implements the following:

$$\forall x_s, \exists n_s^{\text{opt}} : R_s(x_s, n_s^{\text{opt}}) \geq R_s(x_s, n_s), \forall n_s \in [0, |K_s|]. \quad (7)$$

In a such way, the slice tenant will admit a new user only if it generates an increase in revenues, rejecting them otherwise. It is important to remark that

Proposition 1. *The market game \mathcal{G} with the RAC policy of Eq. (7) always admits at least one NE. Furthermore, the Best Response Dynamics (BRD) algorithm always converges to a NE in a finite number of steps.*

Proof. The proposition can be easily verified by construction, by recalling the existence and convergence condition in [12, Theorem 1 and 2]. Indeed, given the formulation in Eq. (7), it can be easily verified that the revenue function $R_s(x_s, n_s^{\text{opt}})$ holds the same property of $R_s(x_s)$ when RAC is not considered, namely being an increasing and differentiable function in x_s , which is a necessary condition for the existence and the convergence to a NE. \square

In what follows, we remove the dependency on n_s of $u_s(\cdot)$ for ease of notation.

3 Theoretical Properties of the Market Game

In this section, we characterize the properties of the market game, based on the customization of the tenants’ private revenue function, $R_s(\cdot)$. The main idea behind our investigation comes from the property that in aggregative games the players can analyze the evolution of the game by simply focusing on their individual action and on the value of the aggregate of the other players. In our specific market structure, we can further simplify those assumptions since players only need to know the updated shape of the pricing function, with no need to expose the actual total resource occupation. This is a very interesting property that allows each player to analyze the game independently on the number of players and on the shape of the opponents’ utility functions, focusing only on the evolution of the prices in the market. Moreover, we can characterize the effects of changing a parameter of the utility function of any player. For this purpose, we introduce the concept of *positive shock* defined in [3].

Definition 1. Consider the payoff function of the game \mathcal{G} , $u_s(x_s, \mathbf{x}_{-s}, p_s)$ with parameter p_s . An increase in $p_s \in \mathbb{R}$ is a positive shock for player s if the payoff function $u_s(\cdot)$ exhibits increasing differences in x_s and p_s and $u_t(\mathbf{x}, p_s) = u_t(\mathbf{x}) \forall t \neq s$.

From [3], we can also state the following

Proposition 2. In the game \mathcal{G} , a positive shock to any player $s \in S$ determines an increase in his allocation, x_s , and decrease in the aggregate allocations of the remaining players, x_{-s} , at the NE. Moreover, it causes an increase in his own payoff and a decrease in at least one other player’s payoff at the NE.

Specifically, we can prove that, for the slicing market game, holds the following

Proposition 3. For any slice s , the admission of new users or an increase in the per-user economic value, χ_k , is a positive shock for the player in the market game \mathcal{G} .

Proof. To prove the proposition, one must verify that the utility function $u_s(\cdot)$ exhibits increasing differences in (x_s, n_s^{opt}) and (x_s, χ_s) , respectively, and that any changing in the value of parameters n_s^{opt} and χ_s for slice s does not affect the utility function of other players $t \in S, t \neq s$. The latter is true by definition, given that the parameters χ_s and n_s^{opt} affect only the revenues, $R_s(x_s)$, which are privately defined by each tenant. The increasing difference property can be verified by checking that the cross-partial derivatives of $u_s(\cdot)$ are nonnegative. This is immediately verified by noticing that $C_s(x_s, \mathbf{x}_{-s})$ does not depend neither on n_s^{opt} nor on χ_s . Therefore, we just need to verify that the cross partials obtained from $R_s(x_s)$ are nonnegative. The analytical derivation is left to the reader. However, intuitively, one can expect that the revenue function satisfies the increasing differences property, being the revenue function $R_s(x_s)$ an increasing function in the purchased resources, x_s , per-user economic value, χ_s , and number of admitted users, n_s^{opt} , as also remarked in Proposition 1. \square

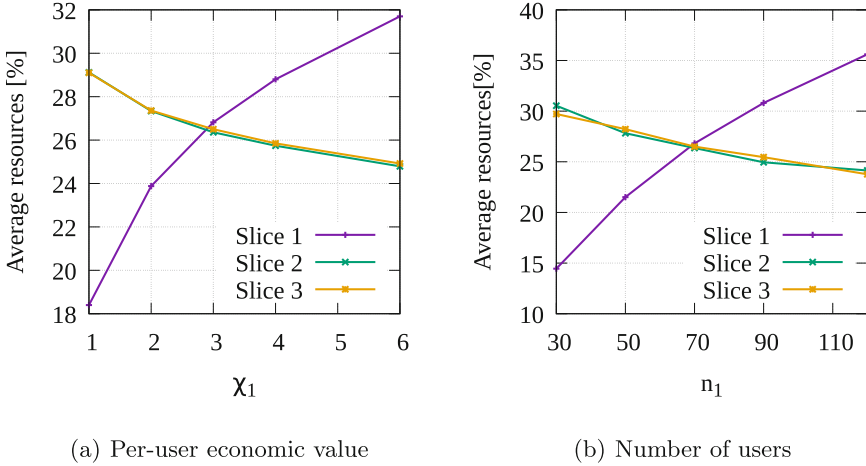


Fig. 1. The parameter χ and number of users cause a positive shock for Slice 1.

An example of a positive shock for a player is plotted in Fig. 1, for the two aforementioned values. We consider a scenario with two symmetric slices and one slice which spans different parameter configurations. In particular, in Fig. 1(a), we show that an increase in the per-user economic value, χ_1 , allows the tenant to get an higher amount of resources, causing a decrease in his opponents' strategies (and, therefore, to the aggregate, as stated in Proposition 2). The same applies when a slice experiences an increase in his traffic demand, expressed as number of accepted users n_1 , where the admission of new users induces the same effect (cf. Fig. 1(b)). Conversely, if we look at the parameters modeling the quality of experience of the users, they do not induce a positive shock in the allocation of the network slices. Indeed, as shown in Fig. 2, an increase in the sensitivity parameter μ and in the maximal throughput r_0 - for example due to a change in the user behavior - may encourage the tenants to purchase more resources in the market (to improve the user performance), but it may also result in the opposite behavior due to non sustainable costs, causing, e.g., an higher user rejection rate and/or lower revenues.

4 Numerical Evaluation

In this section, we simulate the online trading of radio resources in the slicing marketplace. We consider the same slicing setting presented in [12], with $S = \{\text{critical IoT (cIoT), eMBB Premium (eMBB Pr.), eMBB Basic (eMBB Bs.)}\}$, to span different service characteristics and user behaviors, i.e.,

- critical applications with low-rate requirements but high QoS guarantees (cIoT),

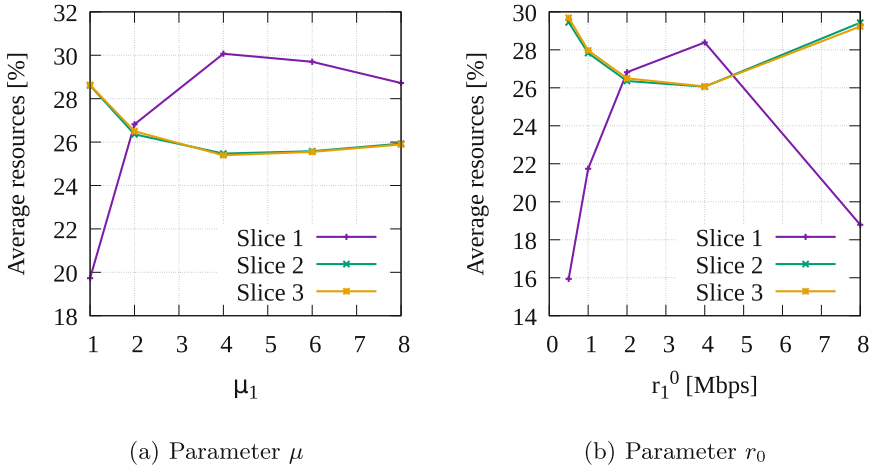


Fig. 2. The parameters μ and r_0 do not cause positive shock for Slice 1.

- non-critical applications with high-rate requirements and high user economic value (eMBB Pr.),
- non-critical applications with adaptive QoS and low user economic value (eMBB Bs.),

that we model by means of different values of the tuple (χ_s, μ_s, r_s^0) . In our experiments, the users of each slice are uniformly distributed throughout the coverage area of 3 cells and their position over time varies according to their random movements. In particular, we assume that the users of the cIoT slice have fixed locations, while the users of eMBB slices move at constant speed of 3 km/h. We further consider, without loss of generality, that all the users within a slice have the same QoS requirements, which are defined in terms of throughput, r_s^0 , and the same economic value, i.e. $\chi_k = \chi_s, \forall k \in K_s$. In Table 1 we resume the parameter settings of the 3 slices.

The simulations reproduce a realistic scenario, where the incoming traffic of the slices vary over time. The simulated traffic traces are taken from real measurements collected from an LTE network in an urban area and they are artificially redistributed among the 3 slices. In our simulations, we consider the traffic patterns of two consecutive weekdays to leverage the network utilization and the resource demand during different time of the day (e.g., showing different behavior during peak hours or during non-peak hours - for example at night). We assume that cIoT end-users have a deterministic network behavior, with constant data transmission during daily hours and only background traffic transmitted during night hours. Contrarily, the eMBB slices exhibit more randomic behavior, due to sudden activation of users requesting data. The traffic traces are taken from an original dataset that consists of time series collecting average downlink datarate observations (measured in bps) with 15-min granularity from 3 cells of an LTE network for a duration of two days. The subdivision of the traffic traces

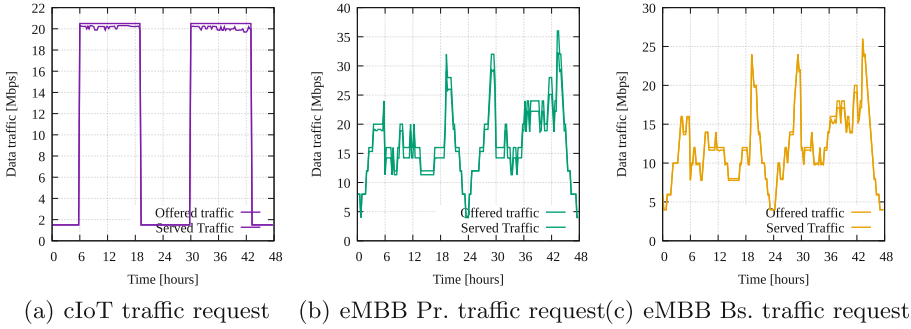


Fig. 3. Time evolution of traffic request for each slice. In dashed lines the incoming offered traffic; in solid lines the accepted served traffic after RAC.

among the 3 slices during the two days of measurements is shown in Fig. 3, for the whole 48 h with 4 samples per hour (for a total of 192 simulated scenarios). For every traffic sample of 15 min, each user random movement is simulated for 3 out of the 15 min to reduce the simulation time, determining the variable serving cell and channel during that time span.

Table 1. Slice parameters setting

Tenant	cIoT	eMBB Pr.	eMBB Bs.
μ_s	8	4	2
r_s^0	0.5 Mbps	4 Mbps	2 Mbps
q_s	0.001	0.001	0.001
χ_s	3	8	3

The experiments are performed in a downlink system level simulator which is 3GPP-calibrated [1] for a 3D Urban Macro (3D-UMa) scenario and abstracts the physical-layer effects through a link-to-system level interface. The interface applies an equivalent Signal-to-Interference-and-Noise Ratio (SINR), computed given the cell topology, the active user transmissions, and a vertically polarized antenna configuration. The radio environment and other relevant simulation parameters are taken from [4]. We assume that the strategy step size of each slice is equal to $\Delta x = 15$ kHz, that is the subcarrier spacing and a total bandwidth of 10 MHz. The traffic traces are then generated in the 3GPP-calibrated system level simulator, where each user data transmission is modeled as Constant Bit Rate (CBR) traffic, and the traffic variation in Fig. 3 is obtained via the activation and/or deactivation of new users in each slice. We assume that the tenants have perfect knowledge of their incoming traffic at time of resource negotiation and they can implement their own independent admission control

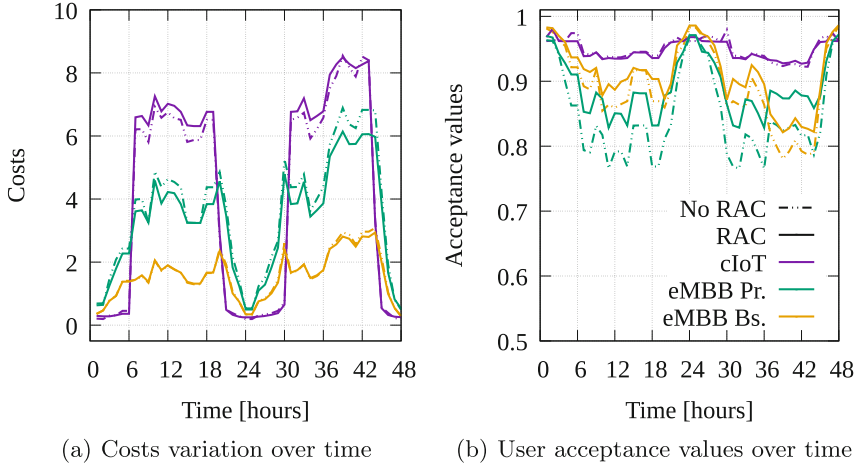


Fig. 4. Real-time KPIs monitoring during marketplace evolution

policies to shape the traffic they are going to accept. The market game is executed at each cell on a timescale of 30 s to let slice tenants adjust the per-cell required resources due to user mobility (e.g., after handovers) and/or variation of channel quality of their users, in addition to traffic fluctuations.

In Fig. 3, we show the offered traffic (in dashed lines) and the served traffic (in solid lines) of each slice, after the RAC policies are implemented. One can notice that both the cIoT and the eMBB Bs. slices (Fig. 3(a) and Fig. 3(c)) can almost serve 100% of their traffic, rejecting only few users during peak hours, while the eMBB Pr. slice shows an higher amount of dropped traffic, specifically in proximity of a burst of traffic for its slice (cf. Fig. 3(v)). This behavior depends on multiple factors, hidden in the strategic behavior of the players in the marketplace, but they can be partly addressed by analyzing the results given in Fig. 4, where we show the achieved performance of the slices with and without RAC. One can observe in Fig. 4(a) the variation in costs for each slice, which follows the increase in price due to their daily traffic oscillation, while in Fig. 4(b) the average acceptance values of the slices, which we use as a measure to estimate the monitored slice KPIs. If we compare the results for the cIoT and the eMBB Bs. when RAC policies are implemented, we see that there is no significant variation in their KPIs. Indeed, given the nature of the two slices, the cIoT slice can always afford higher costs (to accept more users) to keep the desired level of performance, due to the criticality of its service, while the eMBB Bs. slice can always accept users in the system while keeping low costs by offering lower QoS. Slightly variations can be observed only during high traffic peaks, where an optimized utilization of the resources can benefit the slices in terms of lower costs and more affordable performance. Conversely, the eMBB Pr. slice can effectively improve his KPIs only by rejecting part of the traffic, in order to get lower prices and higher performance for the accepted users. However, as shown in the previ-

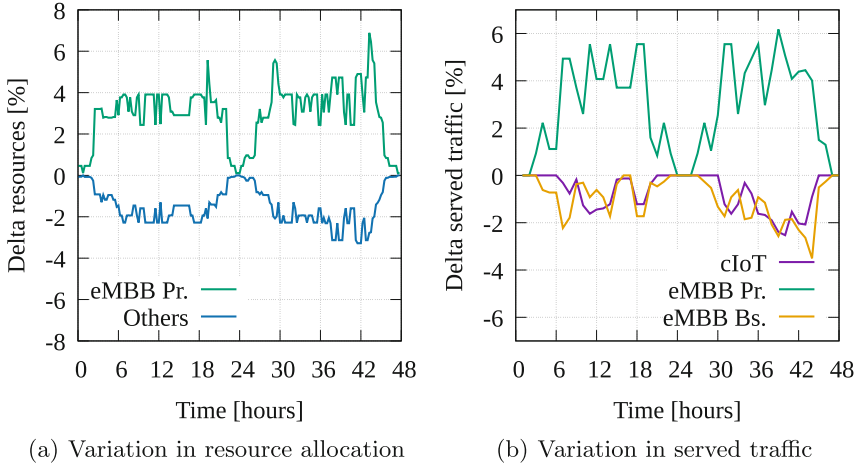


Fig. 5. Positive shock and performance enhancement for eMBB Pr. slice

ous section, the slice tenants can manipulate their configuration parameters to enhance some of their KPIs, being reducing costs or improving performance. In particular, we consider the case when the eMBB Pr. slice intends to improve his performance indicators, namely increasing the percentage of accepted users in the system. For example, by raising the per-user economic value from $\chi_s = 8$ to $\chi_s = 20$, we can observe a radical behavioral change in the marketplace. Indeed, from Fig. 5, one can see the positive shock effect on resource allocation (Fig. 5(a)) and on the served traffic (Fig. 5(b)), plotted as differences with respect to the same values illustrated in previous configuration. In particular, during non-peak hours, where the traffic demand is not high, there are no evident variations in the strategic behavior of all the tenants (due to the lower price for resources); contrarily, the eMBB Pr. can accept more traffic for his slice (by increasing the resource demand) during peak hours. As expected, we can appreciate also the reduction in the resource demand of the other players, with consequences also on their accepted traffic. Although this behavior might look undesirable, it exactly describes how diverse can be the strategic behavior of the slice tenants in this market and how they might influence each other in the decisions to be taken. One can imagine that the optimal setup of the techno-economic KPIs (i.e., the ones reflecting the strategic behavior of the players in the market) can either define the long-term or real-time slice configuration, depending on their private business model. By our approach, we offer slice tenants a mean to optimize the management of their slices, by keeping simultaneously an eye to technical performance indicators, like achieved throughput or latency, and business intents, like cost reductions or profit maximization.

5 Conclusion

In this paper, we discuss a slicing marketplace that enables a dynamic negotiation of radio resources among network slice tenants. We model the interaction of the slice tenants in the market through game theory and describe the theoretical properties of the market game, by means of theoretical demonstrations and numerical simulations. We introduce intra-slice radio admission control policies in the utility function of the players and describe how the tenants behavior in the market is affected by the parameter configuration of the slices. We test the proposed scheme on a dynamic environment, showing that the system automatically scales and adapts the slice resource allocation to the fluctuations of the traffic demand of slice users. Moreover, we show how slice tenants can modify their techno-economic parameter configuration in response to the evolution of the market, e.g., for performance enhancement and/or costs reduction. The automated optimization of those parameters is left to future works.

References

1. 3GPP TR 38.901: Study on channel model for frequencies from 0.5 to 100 GHz (2019)
2. 5GPPP Architecture Working Group: 5G empowering vertical industries (2016). <https://5g-ppp.eu/wp-content/uploads/2016/02/BROCHURE>
3. Acemoglu, D., Jensen, M.K.: Aggregate comparative statics. *Games Econom. Behav.* **81**, 27–49 (2013)
4. Andrews, M., Borst, S., Klein, S., Kroener, H., Mandelli, S.: The effect of additive and multiplicative scheduler weight adjustments on 5G slicing dynamics. In: *IEEE 2nd 5G World Forum (5GWF)* (2019)
5. Badia, L., Lindstrom, M., Zander, J., Zorzi, M.: Demand and pricing effects on the radio resource allocation of multimedia communication systems. In: *IEEE Global Communications Conference (GLOBECOM)* (2003)
6. Bega, D., Gramaglia, M., Fiore, M., Banchs, A., Costa-Perez, X.: DeepCog: cognitive network management in sliced 5G networks with deep learning. In: *IEEE Conference on Computer Communications (INFOCOM)* (2019)
7. Elayoubi, S.E., Jemaa, S.B., Altman, Z., Galindo-Serrano, A.: 5G RAN slicing for verticals: enablers and challenges. *IEEE Commun. Mag.* **57**(1), 28–34 (2019)
8. Jensen, M.: Aggregative games and best-reply potentials. *Econ. Theory* **43**, 45–66 (2010)
9. Jiang, M., Condoluci, M., Mahmoodi, T.: Network slicing in 5G: an auction-based model. In: *IEEE International Conference on Communications (ICC)* (2017)
10. Kelly, F.P., Maulloo, A.K., Tan, D.K.: Rate control for communication networks: shadow prices, proportional fairness and stability. *J. Oper. Res. Soc.* **49**(3), 237–252 (1998)
11. Liang, L., Wu, Y., Feng, G., Jian, X., Jia, Y.: Online auction-based resource allocation for service-oriented network slicing. *IEEE Trans. Veh. Technol.* **68**(8), 8063–8074 (2019)
12. Lieto, A., Malanchini, I., Mandelli, S., Moro, E., Capone, A.: Strategic network slicing management in radio access networks. *IEEE Trans. Mob. Comput.* (2020)

13. Lieto, A., Moro, E., Malanchini, I., Mandelli, S., Capone, A.: Strategies for network slicing negotiation in a dynamic resource market. In: IEEE 20th International Symposium on “A World of Wireless, Mobile and Multimedia Networks” (WoWMoM) (2019)
14. Malanchini, I., Valentin, S., Aydin, O.: Wireless resource sharing for multiple operators: generalization, fairness, and the value of prediction. *Comput. Netw.* **100**, 110–123 (2016)
15. Peha, J.M.: Sharing spectrum through spectrum policy reform and cognitive radio. *Proc. IEEE* **97**, 708–719 (2009)
16. Raza, M.R., Rostami, A., Wosinska, L., Monti, P.: A slice admission policy based on big data analytics for multi-tenant 5G networks. *J. Lightwave Technol.* **37**(7), 1690–1697 (2019)
17. Vassilaras, S., et al.: The algorithmic aspects of network slicing. *IEEE Commun. Mag.* **55**, 112–119 (2017)
18. Wang, G., Feng, G., Tan, W., Qin, S., Wen, R., Sun, S.: Resource allocation for network slices in 5G with network resource pricing. In: IEEE Global Communications Conference (GLOBECOM) (2017)
19. Wang, Q., et al.: Slicenet: end-to-end cognitive network slicing and slice management framework in virtualised multi-domain, multi-tenant 5G networks. In: IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB) (2018)
20. Zheng, J., Caballero, P., de Veciana, G., Baek, S.J., Banchs, A.: Statistical multiplexing and traffic shaping games for network slicing. *IEEE/ACM Trans. Netw.* **26**(6), 2528–2541 (2018)



Service Function Chain Placement in Cloud Data Center Networks: A Cooperative Multi-agent Reinforcement Learning Approach

Lynn Gao¹(✉), Yutian Chen², and Bin Tang³

¹ Data Science, University of California, Irvine, CA 92697, USA
Lmgao@uci.edu

² Economics Department, California State University, Long Beach, CA 90840, USA
Yutian.Chen@csulb.edu

³ Computer Science Department, California State University Dominguez Hills,
Carson, CA 90747, USA
btang@csudh.edu

Abstract. Service function chaining (SFC), consisting of a sequence of virtual network functions (VNFs) (i.e., firewalls and load balancers), is an effective service provision technique in modern data center networks. By requiring cloud user traffic to traverse the VNFs in order, SFC improves the security and performance of the cloud user applications. In this paper, we study how to place an SFC inside a data center to minimize the network traffic of the virtual machine (VM) communication. We take a cooperative multi-agent reinforcement learning approach, wherein multiple agents collaboratively figure out the traffic-efficient route for the VM communication.

Underlying the SFC placement is a fundamental graph-theoretical problem called the k -stroll problem. Given a weighted graph $G(V, E)$, two nodes $s, t \in V$, and an integer k , the k -stroll problem is to find the shortest path from s to t that visits at least k other nodes in the graph. Our work is the first to take a multi-agent learning approach to solve k -stroll problem. We compare our learning algorithm with an optimal and exhaustive algorithm and an existing dynamic programming(DP)-based heuristic algorithm. We show that our learning algorithm, although lacking the complete knowledge of the network assumed by existing research, delivers comparable or even better VM communication time while taking two orders of magnitude of less execution time.

Keywords: Service function chaining · Data centers · Reinforcement learning · k -stroll Problem

1 Introduction

Background and Motivation. Middleboxes (MBs) [9], also known as “network appliances” or “network functions”, are network devices that inspect, filter,

or transform network traffic for purposes of improving network security or performance. Typical examples of MBs include firewalls, intrusion detection systems, load balancers, and network address translators. MBs are an important component of modern enterprise networks to deliver services to user traffic - it has been shown that the number of MBs is comparable with the number of routers in enterprise networks and data centers [32].

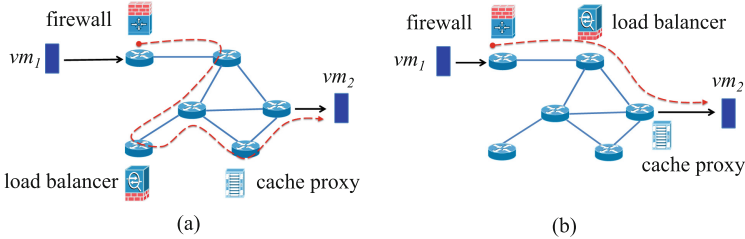


Fig. 1. Illustrating SFC placement in cloud data centers. The SFC consists of three VNFs viz. a firewall, a load balancer, and a cache proxy. The SFC placement in (b) is more network-efficient than the one in (a).

In recent years, with the advancement of *network function virtualization* (NFV) [3], MBs can now be implemented as virtual network functions (VNFs) running as virtual machines (VMs) [29] or containers [33] on commodity hardware platforms. Compared to MBs, which are mostly proprietary and dedicated purpose-built hardware devices, VNFs are software implementations that are cost effective and flexible for deployment. As such, deploying VNFs has become an effective technique in cloud data centers to achieve flexible service management and reduce capital and operational expenditures.

In particular, *service function chaining* (SFC) is established in cloud data centers to require virtual machine (VM) cloud traffic to traverse a chain of VNFs [23, 27, 42]. Figure 1(a) shows an example of SFC, wherein cloud traffic generated at VM vm_1 traverses a sequence of VNFs including a firewall, a load balancer, and a cache proxy, to arrive at VM vm_2 . With this traversal, the SFC blocks malicious traffic detected by the firewall, then diverts the credible VM traffic using the load balancer to avoid network congestion, and finally caches the network packets at the proxy server for quick data access by other cloud users.

SFC Placement Problem. As cloud network resources such as bandwidth and energy are limited in a cloud data center network, one important task for cloud operators is to install the VNFs at the right locations inside the network to optimize the cloud traffic or user-perceived VM communication delay. In this paper, we study the *SFC placement problem*. Given a source and a destination of a VM communication flow inside a cloud data center, and an SFC consisting

of k VNFs of different service functions, it studies where to install the k VNFs inside the network to minimize the communication traffic or delay of the VM flow. In fact, the SFC placement in Fig. 1(a) is not traffic- or delay-optimal for the communicating VM pair (vm_1, vm_2) . By placing three VNFs as shown in Fig. 1(b), the resulting cloud traffic and user-perceived delay are dramatically reduced, as now the traffic only traverses three switches (with two network hops) compared to six switches (with six network hops) in Fig. 1(a).

State-of-the-Art. Tran et al. [36] showed that the SFC placement problem is equivalent to a fundamental graph-theoretical problem called the k -stroll problem [6, 10]. Given a weighted graph $G(V, E)$ and two nodes $s, t \in V$, and an integer k , k -stroll problem is to find a shortest path from s to t that visits at least k other nodes in the graph. As k -stroll is NP-hard, Chaudhuri et al. [10] presented a primal-dual-based $2 + \epsilon$ approximation algorithm. That is, it yields a solution with cost that is at most $2 + \epsilon$ times of the optimal cost. As the primal-dual algorithm is complicated and difficult to implement for a large-scale cloud data center, Tran et al. [36] designed a dynamic programming (DP)-based heuristic algorithm to solve the SFC. They showed that their approach constantly outperforms the performance guarantee of $2 + \epsilon$ provided by Chaudhuri et al. [10]. However, the DP algorithm needs the full knowledge of the network and a complete graph of the network as its input.

Our Contributions. In this paper, we design a cooperative multi-agent reinforcement learning (MARL) algorithm to solve the SFC placement problem. Reinforcement learning (RL) [34] refers to the use of autonomous agents to learn to perform a task by trial and error without human intervention. Unlike traditional computer algorithms, RL executes through the iterative interaction of the agent with the environment to learn about the environment, thus being more adaptive and robust to the dynamic network environment. In addition, as many network-related combinatorial problems are NP-hard and it is time-consuming to find the exact solutions, RL becomes a time-efficient alternative to solve these problems. As such, RL has been well utilized to solve network-related combinatorial optimization problems [2].

However, none of the existing research utilizes RL to solve the k -stroll problem. We show that our MARL algorithm, without knowing the complete graph of the data center network assumed by existing research, delivers comparable or even better VM communication time than existing research, while taking two orders of magnitude less in execution time.

Two characteristics of the k -stroll problem make RL a particularly good candidate to solve the problem. First, to find a shortest route from source s to destination t while visiting at least k other nodes, an agent needs to constantly make decisions along the way. Such multi-step decision making is exactly the kind of problem that RL is designed to solve. Second, the goal of k -stroll is to minimize the travel cost of the agent that constantly makes progress, which corresponds well to the goal of RL of maximizing the cumulative reward.

Paper Organization. The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 formulates the SFC placement problem. In Sect. 4 we propose our MARL algorithm. Section 5 presents the existing DP algorithm as well as an optimal algorithm for SFC placement. Section 6 compares all the algorithms and discusses the results. Section 7 concludes the paper.

2 Related Work

In this section, we first review SFC placement research in general and then SFC placement research using machine learning in particular, to illustrate the contributions of our work.

SFC Placement Research. There has been extensive research for SFC placement, also referred to as VNF placement [25]. Bari et al. [5] studied a VNF orchestration problem that determined the required number and placement of VNFs to optimize network operational costs. It provided an Integer Linear Programming (ILP) solution and an efficient greedy algorithm. Bhamare et al. [7] studied the VNF placement problem that minimized inter-cloud traffic while satisfying deployment cost constraints. They mainly used queueing theories and statistical analysis, which are different from our graph-theoretical approach. Feng et al. [15] studied the NFV service distribution to minimize the overall cloud network resource cost. They formulated a multi-commodity-chain flow problem and provided a $O(\epsilon)$ fast approximation algorithm. Huin et al. [22] studied the SFC placement problem to avoid data passing through unnecessary network devices. Sang et al. [31] and Chen et al. [12] minimized the cost of VNF deployment that provides services to flows and designed optimal and approximate algorithms under different network topologies. However, they did not consider the chain-order sequence of VNFs required in SFCs. Ma et al. [28] considered the traffic changing effect of VNFs and studied the SFC deployment problems with the goal to load-balance the network. Flores et al. [16] studied policy-aware VM migration and placement problem to mitigate the dynamic network traffic in cloud data centers considering that an SFC has already been placed in the network.

There are other works that studied the admission control aspect of SFC placement when not all the user requests can be satisfied due to network resource constraints. They studied how to maximize the total utility [24] and throughput [40] of the satisfied requests, or to maximize the difference between the service provider's profit and the total deployment cost of VNFs [13]. Yang et al. [39] studied how to place VNFs on the edge and public clouds such that the maximum link load ratio is minimized and each user's requested delay is satisfied. Gu et al. [19] designed a dynamic market auction mechanism for the transaction of VNF service chains that achieves near-optimal social welfare in the NFV eco-system.

Unlike most of the above work, we take a graph-theoretical approach and model the SFC placement as a graph-theoretical problem. We uncover that the problem is equivalent to the k -stroll problem. Although k -stroll problem has been studied extensively in the theory community, it has not been studied in a

concrete network context targeting a specific network problem such as the SFC problem in cloud data centers. Due to such theoretical root of our research, the techniques developed in this paper could be utilized in any other network context where the k -stroll model is relevant and applicable.

SFC Placement Research Using Machine Learning. Recently, machine learning techniques have been utilized to solve SFC placement problems. For example, some works employed machine learning to estimate upcoming traffic rates and to adjust VNF deployment [8, 14, 21, 35, 41]. Xiao et al. [38] considered an online SFC placement problem with unpredictable real-time network variations and various service requests and introduced a Markov decision process (MDP) model to capture the dynamic network state transitions. They proposed a deep reinforcement learning (DRL) approach to jointly optimize the operation cost of NFV providers and the total throughput of requests. Using DRL techniques, Pei et al. [30] aimed to minimize the weighted costs of VNF placement cost, VNFI running cost, and penalty of rejected user requests. They proposed a Double Deep Q Network (DDQN)-based optimal solution that places or releases VNF Instances following a threshold-based policy. Recently, Wang et al. [37] extended the above DDQN approach to solving an online fault-tolerant SFC Placement.

All above works adopted the DRL approach. Utilizing neural network-based function approximation algorithms, DRL is a powerful technique that is able to handle complex states and decision-making for agents. As such, DRL is both time- and resource-consuming. Further, unlike DRL that must learn from existing data to train algorithms to find patterns, and then use that to make predictions about new data, RL uses feedback (i.e., rewards) from interacting with the environment to maximize an agent's cumulative reward. Thus the RL model is better suited to solve our SFC placement problem. Besides, as k -stroll problem has a low-dimensional and discrete setting in terms of agent's states and actions, RL is sufficient to solve the SFC placement. We show that the our RL-based solution is competitive to the optimal solution.

Our work is inspired by Ant-Q [17], a family of algorithms that combines RL algorithms with the observation of ant colony behaviors. Ant-Q designed a Q-learning-based algorithm [34] to solve the traveling salesman problem. However, like [36], it assumed that the input is a pre-processed complete graph of the studied network. Our work does not have this assumption and our designed algorithm directly works on any network topologies such as fat-tree data center networks studied in this paper.

3 Problem Formulation

System Model. We model a data center network as an undirected graph $G(V, E)$ where $V = V_h \cup V_s$ includes a set of hosts $V_h = \{h_1, h_2, \dots, h_{|V_h|}\}$ and a set of switches $V_s = \{s_1, s_2, \dots, s_{|V_s|}\}$. E is a set of edges, each connecting either

one switch to another or a switch to a host. Figure 2 shows a $K=4$ fat-tree [4] data center where K is the number of ports per switch.¹

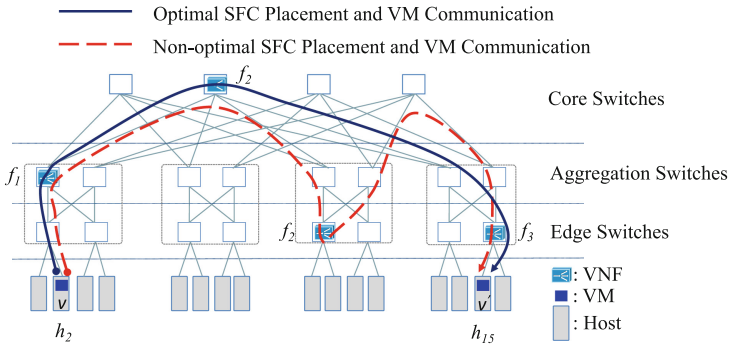


Fig. 2. A K -ary fat-tree data center with $K = 4$. It has 16 hosts: h_1, h_2, \dots , and h_{16} , an SFC with 3 VNFs: f_1, f_2 , and f_3 , and a VM flow: (v, v') . One optimal and one non-optimal SFC placement and communication are shown in the blue solid line and the red dashed line, respectively. (Color figure online)

There is an SFC consisting of n VNFs $\mathcal{F} = \{f_1, f_2, \dots, f_n\}$, each of which needs to be installed (i.e., placed) on a different switch in the data center. Once the VNFs are installed, it requires that the VM traffic to go through the VNFs in the order of f_1, f_2, \dots , and f_n . We refer to f_1 (and f_n) as the *ingress* (and *egress*) VNF, and the switch where the ingress (and egress) VNF is installed as the *ingress* (and *egress*) switch. There is one pair of communicating VMs (v, v') already in the data center, where v is located at host $s(v)$ and v' at $s(v')$. v and v' are referred to as the *source* and *destination VM*, and $s(v)$ and $s(v')$ as *source* and *destination host* respectively.

Each edge $(u, v) \in E$ has a weight $w(u, v)$, indicating either the network delay or energy cost on this edge caused by the VM communication. Given any host or switch u and v , let $c(u, v)$ denote the total cost of all the edges traversed by VM communication from u to v . Thus the *communication cost* of the VM flow (v, v') is $c(s(v), s(v'))$. Note that $c(s(v), s(v'))$ is not necessarily the cost of the shortest path between $s(v)$ and $s(v')$, as the VM communication must traverse a sequence of VNFs.

Figure 2 shows the VM flow (v, v') , where v is located at host h_2 and v' at h_{15} , and one SFC consisting of three VNFs f_1, f_2 and f_3 . The blue solid line shows one optimal SFC placement for this VM flow, which results in 6 hops of VM communication between v and v' . However, if we instead place f_2 at one of the edge switches, the VM communication between v and v' becomes 10 hops, as shown in the red dashed line. Here, we use unweighted costs (i.e., number of

¹ We use fat-trees for illustration purpose. However, the problems and solutions proposed in this paper are applicable to any data center topology.

edges) only for purpose of illustration, as the problem and its solution target weighted graphs. Table 1 shows all the notations.

Table 1. Notation summary

Notation	Description
$G(V, E)$	A data center graph, where $V = V_h \cup V_s$
$w(u, v)$	Weight of an edge $(u, v) \in E$
$c(u, v)$	Cost between hosts (or switches) $u, v \in V$
V_h	$V_h = \{h_1, h_2, \dots, h_{ V_h }\}$ is the set of $ V_h $ hosts
V_s	$V_s = \{s_1, s_2, \dots, s_{ V_s }\}$ is the set of $ V_s $ switches
\mathcal{F}	$\mathcal{F} = \{f_1, f_2, \dots, f_k\}$ is the set of k VNFs of an SFC
(v, v')	The source and destination communicating VMs
$s(v)$	The source host where source VM v is stored
$s(v')$	The destination host where destination VM v' is stored
K	K -ary fat-tree data center wherein each switch has K ports
$p(j)$	SFC placement function p , f_j is placed at switch $p(j) \in V_s$
$C_c(p)$	Total VM communication cost with VNF placement p
α	The learning rate of each agent, $0 \leq \alpha \leq 1$
γ	The discount rate of each agent, $0 \leq \gamma \leq 1$
δ, β	Parameters weighing the relative importance of the Q-value and the edge length in the agent's action selection rule

Problem Formulation. We define a *SFC placement function* as $p : \mathcal{F} \rightarrow V_s$, which places VNF $f_j \in \mathcal{F}$ at switch $p(j) \in V_s$. Given any SFC placement p , denote the communication cost of VM flow v, v' under p as $C_c(p)$. Therefore,

$$C_c(p) = \sum_{j=1}^{n-1} c(p(j), p(j+1)) + \left(c(s(v_i), p(1)) + c(p(n), s(v'_i)) \right). \quad (1)$$

Note that for (v, v') , the ingress switch is always $p(1)$ and the egress switch is always $p(n)$. The objective of the SFC placement problem is to find a p to minimize $C_c(p)$.

k-stroll problem. Previous work [36] has shown that the SFC placement problem is equivalent to *k-stroll problem* [6, 10], which is NP-hard. Given a weighted graph $G=(V, E)$ with nonnegative length w_e on edge $e \in E$, two special nodes s and t , and an integer k , the *k-stroll problem* finds an s - t path or walk (i.e., a stroll) of minimum length that visits at least k distinct nodes excluding s and t . When $s=t$, it is called the *k-tour problem*. The triangle inequality holds for all edges: for $(x, y), (y, z), (z, x) \in E$, $w(x, y) + w(y, z) \geq w(z, x)$.

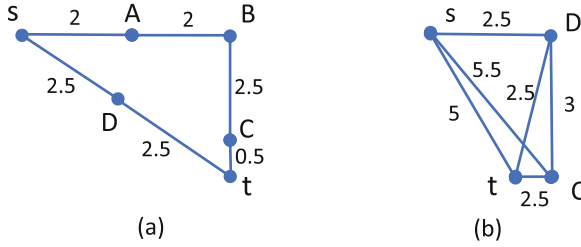


Fig. 3. (a) An example of k -stroll problem. (b) The complete graph for the DP-based Algorithm 2 (only the relevant part for the DP computation is shown).

Figure 3(a) shows a graph of six nodes and six edges with the weight of each edge shown as well. An optimal 2-stroll between s and t is $s, D, t, C,$ and t , with a cost of 6. While most of the works solving k -stroll problem assume the graph is non-complete graph [10, 11, 18], two works [6, 36] assume its input is a complete graph. In particular, Tran et al. [36] showed that by converting the data center graph into a complete graph, it is able to design an efficient dynamic programming (DP) based heuristic. In this paper, however, we show that our reinforcement learning-based algorithm can relax this assumption while still achieving comparable or even better performance compared to the DP-based algorithm. Next we propose a multi-agent reinforcement learning algorithm to solve the SFC placement problem.

4 Multi-agent Reinforcement Learning Algorithm for SFC Placement

In this section, we first present the basics of RL and then our cooperative multi-agent reinforcement learning (MARL) framework for SFC placement.

Reinforcement Learning (RL). In a RL system [34], an agent’s decision making is described by a 4-tuple (S, A, t, r) wherein,

- S is a finite set of states,
- A is a finite set of actions,
- $t : S \times A \rightarrow S$ is a state transition function, and
- $r : S \times A \rightarrow R$ is a reward function, where R is a real value reward.

That is, at a specific state $s \in S$, the agent takes an action $a \in A$ to transition to state $t(s, a) \in S$ while receiving a reward $r(s, a) \in R$. The agent maintains a policy $\pi(s) : S \rightarrow A$ that maps its current state $s \in S$ into the desirable action $a \in A$. We consider a deterministic policy where given the state, the policy outputs a specific action for the agent to take to go to the next step. Deterministic policy suites the SFC placement well as an agent always attempts

to make progress finding a k -stroll from host $s(v)$ to host $s(v')$; that is, it hopes to get closer to $s(v')$ each time it moves to the next node.

A widely used class of RL algorithms are value-based methods [26, 34]. These algorithms try to extract the near-optimal policy based on the value function $V_s^\pi = E\{\sum_{t=0}^{\infty} \gamma^t r(s_t, \pi(s_t)) | s_0 = s\}$, which is the expected value of a discounted future reward sum with the policy π at stage s . Here, γ ($0 \leq \gamma \leq 1$) is a discounted rate and $r(s_t, \pi(s_t))$ is the reward received by the agent at state s_t at time slot t following policy π .

Q-Learning. Q-learning is a family of value-based algorithms [34]. It learns how to optimize the quality of the actions in terms of the Q-value $Q(s, a)$. $Q(s, a)$ is defined as the expected discounted sum of future rewards obtained by taking action a from state s following an optimal policy. The optimal action at any state is the action that gives the maximum Q-value. For an agent at state s , when it takes action a and transitions to the next state t , $Q(s, a)$ is updated as

$$Q(s, a) \leftarrow (1 - \alpha) \cdot Q(s, a) + \alpha \cdot [r(s, a) + \gamma \cdot \max_b Q(t, b)], \quad (2)$$

where α and γ ($0 \leq \alpha, \gamma \leq 1$) are the learning rate and discount rate respectively. In Eq. 2, $r(s, a)$ is the reward obtained if action a is taken at the current state s and $\max_b Q(t, b)$ is the maximum reward that can be obtained from the next state t .

Multi-agent Reinforcement Learning (MARL) Algorithm. In our MARL framework for SFC placement, there are multiple agents that all start from the source host $s(v)$. They work synchronously and cooperatively to learn the state-action Q-table and the reward table and take action accordingly in any of the states. In the context of the SFC placement, the *states* are the nodes (switches or hosts) where agents are located and *actions* are the nodes they move to next. The common task of all the agents is to learn and find a k -stroll: starting from the source host $s(v)$, each visiting at least k other switches, and ending at the destination host $s(v')$. The k distinct switches found in the k -stroll are where the k VNFs will be placed.

One important component of our MARL algorithm is the *action selection rule*; an agent follows such rule to select the node to move to during its k -stroll learning process. It combines exploration and exploitation of an agent; that is, an agent can reinforce the good evaluations it already knows as well as explore new actions. We define the action selection rule in SFC placement as below.

Definition 1. Action Selection Rule of SFC Placement. *The action selection rule specifies, for an agent located at node s , which node t it moves to next. When $q \leq q_0$, where q is a random value in $[0, 1]$ and q_0 ($0 \leq q_0 \leq 1$) is a preset value, it always chooses the node $t = \operatorname{argmax}_{u \in U} \left\{ \frac{[Q(s, u)]^\delta}{[w(s, u)]^\beta} \right\}$ to move to (i.e., the exploitation). Here δ and β are parameters weighing the relative importance of the Q-value and the edge length while U is the set of nodes not visited yet by the agent. Otherwise, the agent chooses a node $t \in U$ to move to by the following distribution: $p(s, t) = \frac{[Q(s, t)]^\delta / [w(s, t)]^\beta}{\sum_{u \in U} [Q(s, u)]^\delta / [w(s, u)]^\beta}$ (i.e., the exploration).*

In the above action selection rule, by exploitation, an agent, located at node s , always moves to a node t that maximizes the learned Q-value weighted by the length of the edge (s, t) . By exploration, it chooses the next node to move to according to the distribution $p(s, t)$, which characterizes how good the nodes are in terms of learned Q-values and the edge lengths. The higher the Q-value and the shorter the edge length, the more desirable the node to move to.

The above action rule is based on ϵ -greedy exploration [34], wherein an agent selects a random action with probability ϵ and selects the best action, which corresponds to the highest Q-value, with probability $1 - \epsilon$. Moreover, our action rule augments ϵ -greedy exploration by taking into account specific features of cloud data center networks (i.e., edge lengths).

MARL Algorithm. Next, we present our MARL algorithm viz. Algorithm 1. There are m agents initially located at the source host $s(v)$ (line 1). Their k -stroll learning takes place in iterations. Each iteration consists of two stages.

The first stage consists of k steps (lines 3–18). In each step, each agent independently takes actions following Definition 1 to move to the next node. At the end of each step, each updates the Q-value of the involved edge. This continues until each agent finds its k -stroll and arrives at destination host $s(v')$. Following [17], each agent maintains a list of visited switches in its memory, so that it knows how to select an unvisited switch to visit while visiting at least k switches before arriving at $s(v')$.

In the second stage (line 19–23), it finds among the m k -strolls the one with the smallest length, and updates the reward value of the edges that belong to this shortest k -stroll as well as the Q-values according to Eq. 2. Finally, it checks if the termination condition is met. If not, it goes to the next iteration and repeats the above two stages. Here, the termination condition is either a specified number of iterations or within some proximity to the costs of the compared DP and optimal algorithms.

Algorithm 1 MARL Algorithm for SFC Placement.

Input: A data center graph $G(V = V_s \cup V_h, E)$, $s(v_1)$, $s(v'_1)$, and an SFC (f_1, f_2, \dots, f_n) .

Output: A k -stroll from $s(v_1)$ to $s(v'_1)$; that is, a switch $p(j) \in V_s$ to place each of the k VNFs $f_j \in \mathcal{F}$ and the cost $C_c(p)$ of the k -stroll.

Notations: i : index for switches; j : index for agents;

U_j : the set of nodes unvisited by agent j , initially $U_j = V_s$, the set of switches;

L_j : the path taken by agent j , initially empty;

l_j : the length of L_j , initially zero;

r_j : the node where agent j is located currently;

$Q(u, v)$: Q-value of edge (u, v) , initially $\frac{|E|}{|V| \cdot \sum_{(u,v) \in E} w(u,v)}$;

p : an array storing the distinct switches on $s(v_1)$ - $s(v'_1)$ stroll;

α : learning rate, $\alpha = 0.1$;

γ : discount factor, $\gamma = 0.3$;

W : a constant value of 10 following [17];

1. Initially all the m agents are at host $s(v)$, i.e., $r_j = s(v), 1 \leq j \leq m$;
2. **while** (termination condition is not met)
3. **for** ($i = 1; i \leq k; i++$) // Finding the k switches to place VNFs
4. **for** ($j = 1; j \leq m; j++$) // Agent j
5. Agent j decides the next node s_j to move to following action rule
 in Definition 1;
6. $L_j = L_j \cup \{s_j\}$;
7. $l_j = l_j + w(r_j, s_j)$;
8. $Q(r_j, s_j) = (1 - \alpha) \cdot Q(r_j, s_j) + \alpha \cdot \gamma \cdot \max_{z \in U_j} Q(s_j, z)$; // Q-value
9. $r_j = s_j$; // Agent j moves to switch s_j ;
10. $U_j = U_j - \{s_j\}$; // Switches not yet visited by agent j
11. **end for**;
12. **end for**;
13. **for** ($j = 1; j \leq m; j++$) // Agent j ends at destination host $s(v')$
14. $L_j = L_j \cup \{s(v')\}$;
15. $l_j = l_j + w(r_j, s(v'))$;
16. $Q(r_j, s_j) = (1 - \alpha) \cdot Q(r_j, s_j) + \alpha \cdot \gamma \cdot \max_{z \in U_k} Q(s_j, z)$; // Q-value
17. $r_j = s(v')$;
18. **end for**;
19. Let $j^* = \operatorname{argmin}_{1 \leq j \leq m} l_j$ be the agent with a k -stroll of smallest length;
20. **for** (each edge $(u, v) \in L_{j^*}$)
21. $r(u, v) = \frac{W}{l_{j^*}}$; // Update reward value $r(u, v)$;
22. $Q(u, v) \leftarrow (1 - \alpha) \cdot Q(u, v) + \alpha \cdot [r(u, v) + \gamma \cdot \max_b Q(v, b)]$; // Q-value
23. **end for**;
24. **end while**;
25. **RETURN** The switch $p(j) \in V_s$ to place VNF $f_j \in \mathcal{F}$ and the cost $C_c(p)$.

Discussion. In each iteration, the first stage takes $m \cdot k$, the second stage takes $m + k$. Assume N iterations take place, then the time complexity of Algorithm 1 is $O(N \cdot m \cdot k)$. A key question for Algorithm 1 is whether it is convergent (i.e., it is able to find the global optimum k -stroll in finite time). Gutjahr [20] gave a graph-based general framework to study convergences of ant systems. It shows that under certain conditions, the solutions generated can converge with a high probability to be arbitrarily close to the optimal solution for a given problem instance. However, as it is a general framework, it does not tackle specific combinatorial problems including the k -stroll. Considering the simple definition and elegant discrete structure inherent in k -stroll problem, studying the convergence as well as estimating the theoretical speed of convergence of applying RL to solve the k -stroll problem is promising future research.

5 Existing Algorithms for SFC Placement

We compare our MARL algorithm with existing work and a naive exhaustive optimal algorithm. We present them below to be self-contained for the paper.

19. **end while;**
20. **if** ($num < k$) $r++$; // less than k distinct switches
21. **else** $found = \text{true}$;
22. **end while;**
23. Place f_1, \dots, f_k on the first k switches stored in p ;
24. **RETURN** $c(s(v), s(v'), r)$.

Example 1. Figure 3(b) shows the complete graph of Fig. 3(a) that is used for above DP computation. In this example, the Algorithm 2 is able to find the optimal 2-stroll between s and t as s, D, t, C , and t .

Optimal Algorithm. Below we present Algorithm 3, an exhaustive algorithm that enumerates all the SFC placements and finds the one with minimum cost, thus solving the SFC placement problem optimally. It takes $O(|V|^k)$. Although it is not time-efficient, it can be implemented easily, and we compare it with other algorithms as a benchmark in small cases.

Algorithm 3 Exhaustive and Optimal SFC Placement.

Input: A data center graph $G(V, E)$, $s(v)$ and $s(v')$, and an SFC (f_1, f_2, \dots, f_k) .

Output: A VNF placement p and the total cost $C_c(p)$.

1. $C_c(p) = +\infty$;
2. Among all $|V_s| \cdot (|V_s| - 1) \cdot \dots \cdot (|V_s| - k + 1)$ SFC placements, find p that gives the minimum cost $C_c(p)$;
3. **RETURN** p and $C_c(p)$.

6 Performance Evaluation

Experiment Setup. We compare our RL-based learning algorithm viz. Algorithm 1 (referred to **RL**) with DP-based algorithm Algorithm 2 (referred to as **DP**) and exhaustive optimal Algorithm 3 (referred to as **Optimal**). We write our own simulator in Python on a MacBook Pro (Big Sur 11.5.1) with Intel Processor (2.7 GHz Quad-Core Intel Core i7) and 16 GB of memory. As Optimal takes a long time to execute, we first compare these three algorithms in small $K = 4$ fat-tree cloud data centers of 16 hosts. We then compare RL with DP in $K = 8$ data centers of 128 hosts. As RL is a multi-agent cooperative learning algorithm, we also investigate the effects of the number of agents m in a $K = 6$ fat-tree data center of 54 hosts. Unless otherwise mentioned, m is set as $\frac{K^3}{4}$, the number of hosts in the data center.

In the plots, each data point is an average of 20 runs with 95% confidence interval. For a fair comparison, in each run instance, the source VM v and destination VM v' are first randomly placed on the hosts then we compare the algorithms on the same VM placement. The SFCs in real-world cases are

broadly categorized into two types viz. access SFCs and application SFCs [1]. As it shows that a typical SFC could have 5 to 6 access functions and 4 to 5 application functions, we consider up to 11 VNFs in an SFC.

Simulation Parameters. Following [17], we set all the RL related parameters as follows. The learning rate $\alpha = 0.1$, the discount factor $\gamma = 0.3$, the original Q-values for all the edges is $\frac{|E|}{|V| \cdot \sum_{(u,v) \in E} w(u,v)}$; that is, 1 divided by the multiplication of average distance between all nodes with number of nodes. For the parameters used in action selection rule, $\delta = 1$ and $\beta = 2$. The constant number W used in Algorithm 1 is set as 10.

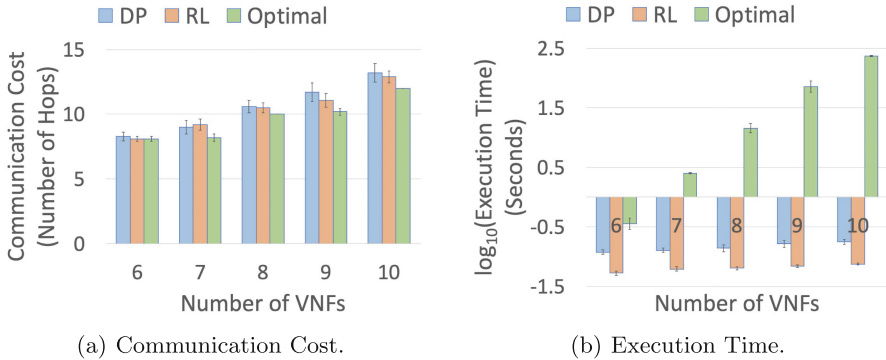


Fig. 4. Comparing RL, DP, and Optimal in $K = 4$ fat-tree data centers. Here, $m = 16$.

Comparing RL, DP, and Optimal in $K = 4$ Fat-Trees. Figure 4 compares all three algorithms in $K = 4$ fat-trees while varying the number of VNFs k . Figure 4(a) compares the VM communication costs in terms of the number of hops yielded by all three algorithms. We have several observations. First, Optimal performs the best by giving the smallest communication cost. Second, RL performs better than DP in most cases although it does not have the complete knowledge of the data center network as DP does. This demonstrates that RL is indeed an effective SFC placement algorithm. Figure 4 show the execution time of the three algorithms. While both RL and DP are time-efficient, incurring less than one second for all the cases, RL takes around half of the time compared to DP (note the logarithmic scale of the y values). In contrast, Optimal takes an enormous amount of time, in the order of hundreds of seconds when k gets large. This shows that our RL algorithm achieves comparable communication cost as the existing SFC placement algorithms while using smaller amount of execution time.

Comparing RL and DP in $K = 8$ Fat-Trees. Next, we compare RL and DP in a larger scale of $K = 8$ fat-tree data centers while varying the number

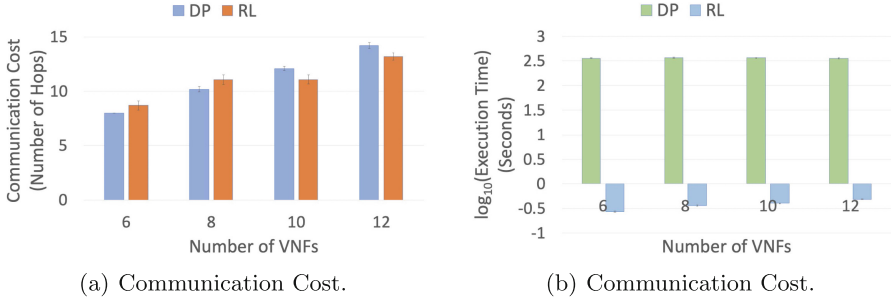


Fig. 5. Comparing RL and DP in $K = 8$ fat-tree data centers, $m = 128$ and $k = 10$.

of VNFs k from 6 to 12. Figure 5(a) shows that DP performs better than RL when k is relatively small and RL performs better than DP when k gets large. However, Fig. 5(b) shows that our RL algorithm has a much smaller amount of execution time than DP (again, note the logarithmic scale of the y values). In particular, while RL takes less than one second to find the SFC placement, the DP takes hundreds of seconds to do so.

This is in sharp contrast to Fig. 4(b), which shows that DP only takes twice as much time as RL does. As the time complexity of DP is $O(k \cdot |V|^4)$ and $|V| = \frac{5}{4} \cdot K^2$, the number of switches in a K -ary fat-tree, the time complexity of DP in a K -ary fat-tree is thus $O(k \cdot K^8)$. On the other hand, the time complexity of RL is $O(N \cdot m \cdot k) = O(N \cdot K^3 \cdot k)$, which is less dependent on the size of the fat-tree than DP is. The two orders of magnitude difference in execution times show that the RL algorithm is a promising technique for the SFC placement problem. Another reason why RL performs better can be attributed to the large number of agents (i.e., 128) that participate in the RL algorithm. As all the agents work cooperatively and synchronously to learn the state-action Q-table and make progress towards finding the k -stroll, the algorithm takes less time.

Effects of Number of Agents m in RL. Finally, we take a close look at the RL algorithm, and investigate the effects of the number of agents m on the performance of RL. We place an SFC of 10 VNFs (i.e., $k = 10$) in a $K = 6$ data center. Table 2 shows that with the increase of m , the VM communication costs found by RL, the execution time, and the number of training iterations for RL all decrease. As more agents participate in the cooperative learning process, its time efficiency increases dramatically. We also observe that the VM communication cost seems to stabilize when m reaches 20. One possible reason could be that when $m = 20$, the RL is able to find the optimal VM communication cost.

Table 2. Varying number of agents m in RL. Here, $k = 10$ and $K = 6$.

Number of agents m	1	5	10	15	20
Communication cost (number of hops)	23.3	17.6	14.7	13.7	13.4
Execution time (seconds)	65.54	27.72	9.07	2.70	0.22
Number of iterations	169.6	75.05	22.25	10.7	3.1

7 Conclusions and Future Work

For data center operators, figuring out how to place SFC with different VNFs inside a data center network while minimizing the communication cost of VMs is a critical task. In this paper, we proposed a multi-agent reinforcement learning algorithm to solve the SFC placement problem. Although the SFC placement problem has been solved extensively with and without resorting to machine learning techniques, our work has two novelties. First, we discovered that at the core of the SFC placement problem is the k -stroll problem, which is only studied in the theory community and has not been identified by any of the existing research on SFC placement. We hope our work can shed some light on how k -stroll can not only advance SFC placement research but also model an even wider range of network applications. Second, all the existing work utilizing machine learning techniques adopts a deep reinforcement learning approach. We showed that under k -stroll modeling of SFC placement, this is not necessary as there are a limited number of states and actions available for the agents. We designed a multi-agent reinforcement learning algorithm where multiple agents cooperatively and synchronously solve k -stroll in order to place SFC. We showed via simulations that our algorithm’s performance is comparable with or even better than the existing algorithms, however, with execution time that is up to two orders of magnitude smaller than that of the existing work.

We have three research directions in the future. First, we will investigate the convergence and convergence speed of our RL algorithm by studying the unique discrete structure of k -stroll. For example, the existing DP-based algorithm provides some insights into how visiting k distinct nodes is related to visiting $k + 1$ distinct edges, which can be solved optimally and efficiently. Second, in our current multi-agent scenario, all the agents cooperate with each other by sharing the same reward and updating the same Q-table. In a rational and game-theoretical environment wherein different agents have different goals and rewards, how agents balance cooperation with competition to find a near-optimal SFC placement efficiently remains largely unexplored. Third, currently, we only consider SFC placement for one VM flow. When there exist multiple VM flows with highly diverse and dynamic traffic rates (e.g., transmission rates and bandwidth demands), how to apply RL techniques to achieve optimal network traffic and communication delay in cloud data centers becomes a challenging problem.

Acknowledgment. This work was supported by NSF Grants CNS-1911191 and CNS-2131309.

References

1. Service function chaining use cases in data centers (IETF). <https://tools.ietf.org/html/draft-ietf-sfc-dc-use-cases-06#section-3.3.1>
2. Mazzyavkina, N., Sviridov, S., Ivanov, S., Burnaev, E.: Reinforcement learning for combinatorial optimization: a survey. *Comput. Oper. Res.* **134**, 105400 (2021)
3. Virtualisation, N.F.: An introduction, benefits, enablers, challenges & call for action. In: White Paper, SDN and OpenFlow World Congress, p. 73 (2012)
4. Al-Fares, M., Loukissas, A., Vahdat, A.: A scalable, commodity data center network architecture. *SIGCOMM Comput. Commun. Rev.* **38**(4), 63–74 (2008)
5. Bari, F., Chowdhury, S.R., Ahmed, R., Boutaba, R., Duarte, O.C.M.B.: Orchestrating virtualized network functions. *IEEE Trans. Netw. Serv. Manag.* **13**(4), 725–739 (2016)
6. Bateni, M.H., Chuzhoy, J.: Approximation algorithms for the directed k-tour and k-stroll problems. In: Serna, M., Shaltiel, R., Jansen, K., Rolim, J. (eds.) APPROX/RANDOM -2010. LNCS, vol. 6302, pp. 25–38. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15369-3_3
7. Bhamare, D., Samaka, M., Erbad, A., Jain, R., Gupta, L., Chan, H.A.: Optimal virtual network function placement in multi-cloud service function chaining architecture. *Comput. Comm.* **102**, 1–16 (2017)
8. Bunyakitanon, M., Vasilakos, X., Nejabati, R., Simeonidou, D.: End-to-end performance-based autonomous VNF placement with adopted reinforcement learning. *IEEE Trans. Cogn. Commun. Netw.* **6**(2), 534–547 (2020)
9. Carpenter, B., Brim, S.: Middleboxes: taxonomy and issues (2002). <https://tools.ietf.org/html/rfc3234>
10. Chaudhuri, K., Godfrey, B., Rao, S., Talwar, K.: Paths, trees, and minimum latency tours. In: 44th Annual IEEE Symposium on Foundations of Computer Science, pp. 36–45. IEEE (2003)
11. Chekuri, C., Korula, N., Pál, M.: Improved algorithms for orienteering and related problems. *ACM Trans. Algorithms* **8**(3), 1–27 (2012)
12. Chen, Y., Wu, J., Ji, B.: Virtual network function deployment in tree-structured networks. In: 2018 IEEE 26th International Conference on Network Protocols (ICNP), pp. 132–142. IEEE (2018)
13. Eramo, V., Miucci, E., Ammar, M., Lavacca, F.G.: An approach for service function chain routing and virtual function network instance migration in network function virtualization architectures. *IEEE ACM Trans. Netw.* **25**(4), 2008–2025 (2017)
14. Fei, X., Liu, F., Xu, H., Jin, H.: Adaptive VNF scaling and flow routing with proactive demand prediction. In: IEEE INFOCOM 2018-IEEE Conference on Computer Communications, pp. 486–494. IEEE (2018)
15. Feng, H., Llorca, J., Tulino, A.M., Raz, D., Molisch, A.F.: Approximation algorithms for the NFV service distribution problem. In: IEEE INFOCOM 2017-IEEE Conference on Computer Communications, pp. 1–9. IEEE (2017)
16. Flores, H., Tran, V., Tang, B.: PAM & PAL: policy-aware virtual machine migration and placement in dynamic cloud data centers. In: IEEE INFOCOM 2020-IEEE Conference on Computer Communications, pp. 2549–2558. IEEE (2020)

17. Gambardella, L.M., Dorigo, M.: Ant-Q: a reinforcement learning approach to the traveling salesman problem. In: Machine learning proceedings 1995, pp. 252–260. Morgan Kaufmann (1995)
18. Garg, N.: Saving an ϵ : a 2-approximation for the k-MST problem in graphs. In: Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, pp. 396–402 (2005)
19. Gu, S., Li, Z., Wu, C., Huang, C.: An efficient auction mechanism for service chains in the NFV market. In: IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, pp. 1–9. IEEE (2016)
20. Gutjahr, W.: A graph-based ant system and its convergence. *Futur. Gener. Comput. Syst.* **16**, 873–888 (2000)
21. Huang, X., et al.: Online VNF chaining and predictive scheduling: optimality and trade-offs. *IEEE/ACM Trans. Netw.* **29**(4), 1867–1880 (2021)
22. Huin, N., Jaumard, B., Giroire, F.: Optimal network service chain provisioning. *IEEE/ACM Trans. Netw.* **26**(3), 1320–1333 (2018)
23. Joseph, D.A., Tavakoli, A., Stoica, I.: A policy-aware switching layer for data centers. In: Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication, pp. 51–62 (2008)
24. Kuo, T., Liou, B., Lin, K.C., Tsai, M.: Deploying chains of virtual network functions: on the relation between link and server usage. *IEEE/ACM Trans. Netw.* **26**(4), 1562–1576 (2018)
25. Laghrissi, A., Taleb, T.: A survey on the placement of virtual resources and virtual network functions. *IEEE Commun. Surv. Tutor.* **21**(2), 1409–1434 (2019)
26. Littman, M.L.: Value-function reinforcement learning in Markov games. *Cogn. Syst. Res.* **2**(1), 55–66 (2001)
27. Liu, J., Li, Y., Zhang, Y., Su, L., Jin, D.: Improve service chaining performance with optimized middlebox placement. *IEEE Trans. Serv. Comput.* **10**(4), 560–573 (2017)
28. Ma, W., Beltran, J., Pan, D., Pissinou, N.: Traffic aware placement of interdependent NFV middleboxes. *IEEE Trans. Netw. Serv. Manage.* **16**(4), 1303–1317 (2019)
29. Mijumbi, R., Serrat, J., Gorricho, J.L., Bouten, N., Turck, F.D., Boutaba, R.: Network function virtualization: state-of-the-art and research challenges. *IEEE Commun. Surv. Tutor.* **18**(1), 236–262 (2015)
30. Pei, J., Hong, P., Pan, M., Liu, J., Zhou, J.: Optimal VNF placement via deep reinforcement learning in SDN/NFV-enabled networks. *IEEE J. Sel. Areas Commun.* **38**(2), 263–278 (2020)
31. Sang, Y., Ji, B., Gupta, G.R., Du, X., Ye, L.: Provably efficient algorithms for joint placement and allocation of virtual network functions. In: IEEE INFOCOM 2017-IEEE Conference on Computer Communications, pp. 1–9. IEEE (2017)
32. Sherry, J., Hasan, S., Scott, C., Krishnamurthy, A., Ratnasamy, S., Sekar, V.: Making middleboxes someone else’s problem: network processing as a cloud service. In: Proceedings of the ACM SIGCOMM (2012)
33. Soltesz, S., Pötzl, H., Fiuczynski, M.E., Bavier, A., Peterson, L.: Container-based operating system virtualization: a scalable, high-performance alternative to hypervisors. *SIGOPS Oper. Syst. Rev.* **41**(3), 275–287 (2007)
34. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. The MIT Press, Cambridge (2020)
35. Tang, L., He, X., Zhao, P., Zhao, G., Zhou, Y., Chen, Q.: Virtual network function migration based on dynamic resource requirements prediction. *IEEE Access* **7**, 112348–112362 (2019)

36. Tran, V., Sun, J., Tang, B., Pan, D.: Traffic-optimal virtual network function placement and migration in dynamic cloud data centers. Submitted to IEEE IPDPS (2022)
37. Wang, L., Mao, W., Zhao, J., Xu, Y.: DDQP: a double deep Q-learning approach to online fault-tolerant SFC placement. *IEEE Trans. Netw. Serv. Manage.* **18**(1), 118–132 (2021)
38. Xiao, Y., et al.: NFVdeep: adaptive online service function chain deployment with deep reinforcement learning. In: *Proceedings of the International Symposium on Quality of Service*, pp. 1–10 (2019)
39. Yang, S., Li, F., Trajanovski, S., Chen, X., Wang, Y., Fu, X.: Delay-aware virtual network function placement and routing in edge clouds. *IEEE Trans. Mob. Comput.* **20**(2), 445–459 (2019)
40. Zhang, Q., Liu, F., Zeng, C.: Adaptive interference-aware VNF placement for service-customized 5G network slices. In: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 2449–2457. IEEE (2019)
41. Zhang, X., Wu, C., Li, Z., Lau, F.C.: Proactive VNF provisioning with multi-timescale cloud resources: fusing online learning and online optimization. In: *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9. IEEE (2017)
42. Zhang, Y., et al.: Steering: a software-defined networking for inline service chaining. In: *2013 21st IEEE International Conference on Network Protocols (ICNP)*, pp. 1–10. IEEE (2013)



The Art of Concession in General Lotto Games

Rahul Chandan¹(✉), Keith Paarporn¹, Dan Kovenock², Mahnoosh Alizadeh¹,
and Jason R. Marden¹

¹ University of California, Santa Barbara, Santa Barbara CA 93106, USA
{rchandan,kpaarporn,alizadeh,jrmarden}@ucsb.edu

² Chapman University, One University Drive, Orange, CA 92866, USA
kovenock@chapman.edu

Abstract. Success in adversarial environments often requires investment into additional resources in order to improve one’s competitive position. But, can intentionally decreasing one’s own competitiveness ever provide strategic benefits in such settings? In this paper, we focus on characterizing the role of “concessions” as a component of strategic decision making. Specifically, we investigate whether a player can gain an advantage by either conceding budgetary resources or conceding valuable prizes to an opponent. While one might naïvely assume that the player cannot, our work demonstrates that – perhaps surprisingly – concessions do offer strategic benefits when made correctly. In the context of General Lotto games, we first show that neither budgetary concessions nor value concessions can be advantageous to either player in a 1-vs.-1 scenario. However, in settings where two players compete against a common adversary, we find opportunities for one of the two players to improve her payoff by conceding a prize to the adversary. We provide a set of sufficient conditions under which such concessions exist.

Keywords: Game theory · Resource allocation · General lotto games

1 Introduction

Strategic advantages are often held by competitors that possess more budgetary resources that can be invested in more advanced technology, research, or surveillance in order to improve one’s competitive position against opponents. Such factors are central to many domains that feature competitive interactions, such as airport security [20,27], wildlife protection [30], market economics [18], and political campaigning [26]. In this paper, we analyze “concessions” as a viable, alternative component of strategic decision-making in adversarial environments. In particular, we seek to identify whether or not conceding one’s competitive

This work is supported by UCOP Grant LFR-18-548175, ONR grant #N00014-20-1-2359, AFOSR grants #FA9550-20-1-0054 and #FA9550-21-1-0203, and the Army Research Lab through the ARL DCIST CRA #W911NF-17-2-0181.

position can ever be advantageous. Intuitively, concessions would appear to be contradictory to the conventional wisdom on how to gain a strategic advantage, e.g., investing in more resources or information, as concessions weaken one’s competitive position. Nonetheless, this paper demonstrates that such intuition is false as appropriately chosen concessions can often be strategically beneficial.

Within the framework of General Lotto games, we study two types of concessions. The first type, which we term *budgetary concessions*, involves willingly reducing one’s resource budget. The act of “money burning” serves as an analogy for this type of concession. The second type of concession, which we term *battlefield concessions*, involves voluntary non-participation on a non-zero valued battlefield. An appropriate analogy for this type of concession from economics is “market abandonment”. In these scenarios, we assume that concessions are announced to all other players immediately after they are made, such that the other players can respond strategically to the modified competitive environment.

General Lotto games, Colonel Blotto games, and other contest models offer a flexible framework to generate basic insights about the interplay between a competitor’s performance guarantees and the amount of resources reserved for competition [3, 6, 9, 13, 14, 23]. In common formulations, two opposing players have limited resource budgets to allocate to multiple battlefields. A player wins a battlefield and its associated value if she sends more resources than her opponent. To study the role of concessions as a strategic component, we continue this section with a brief overview of General Lotto games and describe our extensions that allow us to study concessions under this model. We also provide a summary of our contributions, namely, the identification of settings where concessions are beneficial. Finally, we draw connections between our work and the related literature.

1.1 General Lotto Games with Concessions

The General Lotto game is played between two opposing players, A and B , who each have a limited budget of resources $X_A, X_B \geq 0$. The players compete over a set of n battlefields $\mathcal{B} = \{1, \dots, n\}$, where a player wins a battlefield $b \in \mathcal{B}$ and its value $v_b \geq 0$ by allocating more resources to b than the opponent. The players make moves simultaneously (i.e., a one-shot game). Each player can use randomized allocations such that the resources spent do not exceed its limited budget in expectation. We denote an instance of the General Lotto game with $\text{GL}(X_A, X_B, \mathbf{v})$, where $\mathbf{v} \in \mathbb{R}_{\geq 0}^n$ is the vector of battlefield valuations. The equilibrium strategies and payoffs in any instance are characterized in the existing literature [13, 15], and we reproduce these in Sect. 2.

We consider the following extension in order to study the strategic role of concessions in General Lotto games: One of the players, say player B , has the option to either voluntarily reduce her own resource budget, or to voluntarily withdraw completely from a chosen battlefield, before engaging with A in the resulting General Lotto game. Specifically, B selects one of the following options:

- *Budgetary concession*: Player B selects some nonzero value $x \in [0, X_B]$, whereupon her resource budget is reduced from X_B to $X_B - x$.

- *Battlefield concession*: Player B selects a battlefield $b \in \mathcal{B}$. The value of the battlefield, v_b , is immediately awarded to player A .

The complete competitive interaction between A and B occurs in two stages. In Stage 0, B decides to concede either budgetary resources or a battlefield to A , as described above. Player B 's decision in this stage then becomes binding and common knowledge. Subsequently, in Stage 1, the players engage in the resulting General Lotto game. If a budgetary concession of $x \in [0, X_B]$ was made in Stage 0, the game $GL(X_A, X_B - x, \mathbf{v})$ is played and the players receive their respective equilibrium payoffs. If a battlefield concession of $b \in \mathcal{B}$ was made in Stage 0, the value v_b is immediately awarded to player A , and the game $GL(X_A, X_B, \mathbf{v}_{-b})$ is played. Here, \mathbf{v}_{-b} is the vector of valuations for the battlefields $\mathcal{B} \setminus \{b\}$. We say that a player has a *beneficial concession* if there exists any concession such that the player secures a strictly higher payoff than her payoff in the nominal General Lotto game (i.e. without concessions). For example, if player B has a beneficial budgetary concession in the General Lotto game, then there exist parameters $X_A, X_B > 0$, $\mathbf{v} \in \mathbb{R}_{\geq 0}^n$ and $x \in [0, X_B]$ such that B 's equilibrium payoff is greater in $GL(X_B - x, X_A, \mathbf{v})$ than in $GL(X_B, X_A, \mathbf{v})$. Our first contribution is as follows:

Contribution 1. *There never exist concessions of either type that improve a player's payoff in the General Lotto game (Proposition 1).*

1.2 Three-Player General Lotto Games with Concessions

Contribution #1 conforms with the conventional intuition that concessions only ever weaken one's position in competitive scenarios. We thus seek to address whether this phenomenon holds more generally. To that end, we shift our focus to a three-player setting, in which players B and C compete in General Lotto games against a common adversary A over two disjoint sets of battlefields $\mathcal{B}_B, \mathcal{B}_C$ whose valuations are given by the vectors $\mathbf{v}_B, \mathbf{v}_C$, respectively. This formulation was first proposed and studied in [16]. The top diagram in Fig. 1a depicts a *nominal* three-player Lotto game (under no concession options). We consider the case where only player B has the option to make concessions. The competitive interaction occurs over three stages as follows, where players' actions become binding and common knowledge in subsequent stages:

- *Stage 0*: Player B decides to make either a budgetary or battlefield concession;
- *Stage 1*: Player A deploys resources $X_{A,B}, X_{A,C} \geq 0$ to the two competitions against B and C , where $X_{A,B} + X_{A,C} \leq X_A$ must be satisfied; and,
- *Stage 2*: Player A engages in the two resulting General Lotto games. If a budgetary concession of $x \in [0, X_B]$ was made in Stage 0, then she plays the game $GL(X_{A,B}, X_B - x, \mathbf{v}_B)$ against player B . Else, if a battlefield concession of $b \in \mathcal{B}_B$ was made in Stage 0, then she plays $GL(X_{A,B}, X_B, \mathbf{v}_{B,-b})$ against B , where $\mathbf{v}_{B,-b}$ denotes the vector of valuations for battlefields $\mathcal{B} \setminus \{b\}$. The game $GL(X_{A,C}, X_C, \mathbf{v}_C)$ is played against player C .

The bottom diagram in Fig. 1a depicts the scenario following a battlefield concession. In Stage 1, we assume player A employs an optimal division of resources such that her cumulative payoff from the two General Lotto games in Stage 2 is maximized. Such optimal divisions are characterized in the literature by [16], and we reproduce these results in the forthcoming Sect. 2. In this three-player setting, we say that B has a *beneficial concession* if there exist any concessions such that B secures a strictly higher payoff than her payoff in the nominal three-player General Lotto game, i.e. if B were to make no concession in Stage 0. Our second contribution is as follows:

Contribution 2. *In three-player General Lotto games, there never exist budgetary concessions that improve a player’s payoff (Theorem 1); however, there do exist battlefield concessions that can improve a player’s payoff. Theorem 2 provides a set of sufficient conditions for when such opportunities are available.*

In the standard, two-player General Lotto game, we show that beneficial concessions do not exist, and, indeed, our result concerning budgetary concessions in the three-player General Lotto game further supports this naïve intuition. However, our results show that beneficial battlefield concessions do exist in three-player General Lotto games, contradicting the conventional wisdom on what constitute viable mechanisms for gaining strategic advantages. More generally, our results suggest that concessions do, in fact, represent reasonable strategic options in competitive interactions.

As our main objective is to establish whether it is possible that beneficial concessions exist in three-player General Lotto games, we consider budgetary concessions of value $x \in [0, X_B]$ and show that $x = 0$ maximizes player B ’s final payoff. Similarly, in the case of battlefield concessions, instead of considering possible values in the vector \mathbf{v}_B , we focus on identifying the battlefield concession value $v \in [0, \Phi_B]$ that maximizes player B ’s final payoff, where Φ_B is the cumulative value of battlefields in \mathcal{B}_B . When $v > 0$, B has a beneficial battlefield concession in any corresponding three-player General Lotto game with $v_b = v$ for some battlefield $b \in \mathcal{B}_B$. In the following example, we identify the occurrence of beneficial battlefield concessions for player B , and the magnitude of B ’s payoff improvement under various parameterizations of the three-player General Lotto game:

Consider a three-player General Lotto game in which players’ initial budget endowments satisfy $X_A, X_B \in [0, 4]$ and $X_C = 1$, and where the cumulative battlefield values in fronts \mathcal{B}_B and \mathcal{B}_C are $\Phi_B = 1.5$ and $\Phi_C = 1$, respectively. For every such game, we compare player B ’s payoff in the nominal game against her payoff after a battlefield concession of each value $v \in [0, \Phi_B]$, and identify the optimal battlefield concession value v^{opt} . Figure 1b illustrates the regime of initial player budgets in which there exist battlefield concessions of any value $v \in [0, \Phi_B]$ such that player B ’s payoff in the resulting game is strictly higher than her payoff in the corresponding nominal game (i.e., the regime where there exists a beneficial battlefield concession for B). Figure 1c shows the percentage improvement over player B ’s payoff in the nominal game associated with

conceding the corresponding optimal battlefield concession value v^{opt} . We plot this percentage improvement for $X_A \in [0, 2.5]$, and $X_B = 0.5$ (dashed line) or $X_B = 1.25$ (solid line).

Intuitively, our results illustrate that battlefield concessions in the three-player General Lotto game – if done properly – can redirect more of player A 's budget toward player C 's set of battlefields, rather than drawing more of A 's budget to B 's remaining set of battlefields, as the remaining value on B 's set of battlefields, $\Phi_B - v^{\text{opt}}$, becomes less of a priority for A . In a sense, the conceding player “appeases” the common adversary by freely offering up a portion of the cumulative battlefield value, and faces less competition as a result. The presence of the additional player C is critical for there to be benefits derived from such concessions. In contrast, budgetary concessions invite A to further pursue her contest against the weakened player B . A budgetary concession reduces B 's strength with no change to the cumulative value of the battlefields. This increases the ratio between value and strength on \mathcal{B}_B , and leads to A seeking even more value from that front.

1.3 Related Works

A primary line of research in Colonel Blotto games focuses on characterizing its equilibria. Since Borel's initial study [3], many works have advanced this thread over the last one hundred years [2, 9, 15, 17, 23, 24, 28]. However, solutions to the most general settings remain as open problems. As such, there are several variants of the Colonel Blotto game that have been studied extensively, none more so than the General Lotto game [1, 13, 15, 19]. Notably, the players' equilibrium payoffs in the General Lotto games have been fully characterized [13, 15]. Due to its tractability, the General Lotto game is often adopted in studies of more complex adversarial environments, including engineering domains such as network security [7, 10, 25] and the security of cyber-physical systems [5, 12].

Our work in this paper is closest to a recent thread in the literature on similar sequential Colonel Blotto and General Lotto games, where players have the option to publicly announce their strategic intentions ahead of play. The three-player General Lotto game was first introduced in [16], who study their own variant model where in Stage 0, players B and C have the opportunity to form an alliance that takes the form of a unilateral budgetary transfer between the players. It is shown that there are cases in which the two players can make unilateral budgetary transfers that are mutually beneficial. Subsequent work in [11, 12] considers similar settings where the two players can decide to add battlefields in addition to transferring resources amongst each other. Counter-intuitively, under this model, both the players achieve better payoff if the transfers are publicly announced to their adversary. The authors of [4] identify a sufficient condition for when publicly pre-committing resources to battlefields offers strategic advantages in the same three player setting. Pre-commitments are a broader class of concessions, where instead of giving away value, the pre-committing player puts a price in terms of budgetary resources on a battlefield.

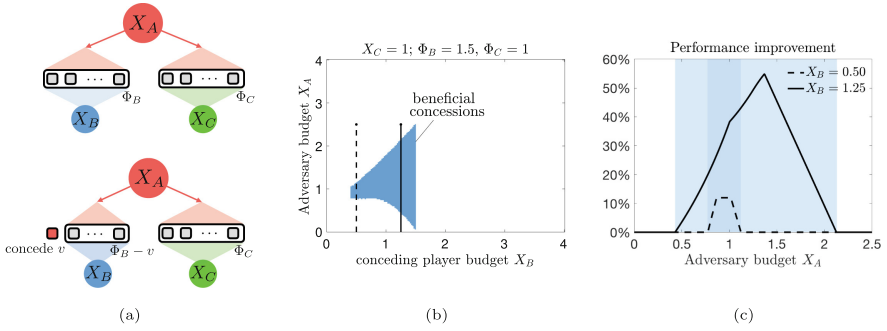


Fig. 1. (a) The top diagram depicts the nominal three player General Lotto game, where the adversary (A) must decide how to divide its endowment to two separate fronts of battlefields, with cumulative values of Φ_B and Φ_C , respectively. The optimal division for A and the resulting payoffs are well-known from the literature [16]. The bottom diagram shows a scenario where player B concedes a battlefield of value $v \in [0, \Phi_B]$. The adversary responds by re-calculating her optimal division based on the modified environment. We seek to answer whether B can benefit from concessions. (b) The parameter region (in blue) where player B has an incentive to concede battlefields. Here, we set $X_C = 1$ and the total valuations of the two fronts are $\Phi_B = 1.5$, $\Phi_C = 1$. (c) The percentage improvement over player B 's payoff in the nominal three player game (without concessions) associated with the optimal battlefield concession. We plot the improvements when $X_B = 0.50$ (dashed line) and $X_B = 1.25$ (solid line) for all values $X_A \in [0, 2.5]$, as depicted in (b). (Color figure online)

The pre-commitment of resources is also studied in [29], but in a different context that involves favouritism. In that work, a one-shot Colonel Blotto game is studied where resources are pre-allocated non-strategically over the various battlefields. More broadly, interest in the role of pre-emption and information in contests has popularized the analysis of sequential versions of these game models in which one player leads the strategic interaction and the other follows (see, e.g., [8, 21, 22]). Though the selection of concessions under our proposed framework is also sequential, we note that the players' strategic interactions still occur simultaneously in the final stage.

The formation of alliances such as those studied in [11, 12, 16] is often not possible, either because mechanisms for coordination between the agencies are not available or because the agencies' budgets are not directly transferable. In contrast, concessions offer a means for a player to improve her competitive position, even when mutual coordination is not possible. Another notable difference between concessions and alliances is that, while alliances can only lead to mutually beneficial outcomes for the players involved, our results suggest that any benefits derived from concessions by one player must come at an expense to the other.

2 Model

In this section, we review useful background on the standard, two-player General Lotto game, then formalize the three-player General Lotto game model.

2.1 Background on General Lotto Games

The standard General Lotto game consists of two players A and B with respective, fixed budgets $X_A, X_B > 0$ competing over the set of n battlefields $\mathcal{B} = \{1, \dots, n\}$ (i.e., front). A player wins on a battlefield b by allocating more budget to b than her opponent, and otherwise loses on b .¹ For each battlefield $b \in \mathcal{B}$, the winning player receives her value $v_b \geq 0$, while the losing player receives zero. Let $\mathbf{v} \in \mathbb{R}_{\geq 0}^n$ denote the vector of battlefield values. An allocation is any vector $\mathbf{x} \in \mathbb{R}_{\geq 0}^n$, where x_b denotes the amount of budget allocated to battlefield b . An admissible strategy for each player $i \in \{A, B\}$ is an n -variate distribution F_i on $\mathbb{R}_{\geq 0}^n$ that satisfies the following budget constraint:

$$\mathbb{E}_{\mathbf{x} \sim F_i} \left[\sum_{b \in \mathcal{B}} x_b \right] \leq X_i. \tag{1}$$

Intuitively, a player may select any distribution over vectors $\mathbf{x} \in \mathbb{R}_{\geq 0}^n$ such that the budget expenditure does not exceed her budget *in expectation*. Each player aims to maximize the expected value won over the battlefields. We observe that the game is a two-player, constant-sum game played in a single stage (Stage 1), and that an instance of the game can be succinctly denoted as $\text{GL}(X_A, X_B, \mathbf{v})$. The General Lotto game is a relaxation of the Colonel Blotto game [3], in which the players' allocations must satisfy the budget constraint with probability 1.

The equilibrium characterization of the General Lotto game is well-understood [13, 15], and each instance $\text{GL}(X_A, X_B, \mathbf{v})$ is known to admit unique equilibrium payoffs as follows:

Fact 1. Let $\text{GL}(X_A, X_B, \mathbf{v})$ denote an instance of the General Lotto game, and $\Phi = \sum_{b \in \mathcal{B}} v_b$. The equilibrium payoff to player $i \in \{A, B\}$ is $\Phi \cdot L(X_i, X_{-i})$, where

$$L(X_i, X_{-i}) = \begin{cases} \frac{X_i}{2X_{-i}} & \text{if } X_i \leq X_{-i} \\ 1 - \frac{X_{-i}}{2X_i} & \text{if } X_i > X_{-i}, \end{cases} \tag{2}$$

and $-i \in \{A, B\} \setminus \{i\}$ is the opposing player.

As discussed in Sect. 1.1, concessions in the two player General Lotto game can be considered by introducing an additional stage (Stage 0) that occurs before the players engage in the General Lotto game (Stage 1). Recall that, in Stage 0,

¹ In the case that the players allocate the same amount of budget to a battlefield, the player with higher overall budget is conventionally awarded the win. However, the choice of tie-breaking rule has no effect on equilibrium characterizations of General Lotto games [15], and hence, our results.

player B makes either a budgetary concession or battlefield concession, which then becomes binding and common knowledge before Stage 1 is played. In the following proposition, we show that neither type of concession can ever increase a player’s payoff over her payoff in the nominal General Lotto game:

Proposition 1. *Consider the General Lotto game with $X_A, X_B \geq 0$ and $\Phi \geq 0$. Neither player can benefit from either a budgetary or battlefield concession.*

Proof. We consider the scenario where player B makes either a budgetary or battlefield concession in Stage 0. Since we make no assumption on the players’ relative strengths, considering player B ’s perspective is without loss of generality.

Firstly, from the equilibrium payoffs identified in Fact 1, if player B makes a budgetary concession, i.e., $X'_B \leq X_B$, then it follows that $\Phi \cdot L(X'_B, X_A) \leq \Phi \cdot L(X_B, X_A)$ since, for fixed y , $L(x, y)$ is monotonically increasing in x . Second, and finally, if player B makes a battlefield concession, i.e., $\Phi' \leq \Phi$, then $\Phi' \cdot L(X_B, X_A) \leq \Phi \cdot L(X_B, X_A)$ since L is nonnegative. \square

2.2 Three-Player General Lotto Games with Concessions

We have shown that concessions cannot provide payoff improvements in the two-player General Lotto game. Thus, we consider the three-player game model proposed in [16] for the remainder of this manuscript. This game consists of players A, B and C with respective budgets $X_A, X_B, X_C > 0$. Player A is engaged in simultaneous General Lotto games against the players B and C over the respective, disjoint fronts \mathcal{B}_B and \mathcal{B}_C . The game is played in two stages: in Stage 1, player A allocates her budget between the two fronts; and, in Stage 2, the two resulting General Lotto games are played. In Stage 2, players B and C receive the payoffs from their respective General Lotto games, and player A receives the sum of her expected payoffs from both General Lotto games. An instance of the game can be succinctly denoted as $3GL(X_A, X_B, X_C, \mathbf{v}_B, \mathbf{v}_C)$, where \mathbf{v}_i denotes the vector of battlefield values in front \mathcal{B}_i , $i \in \{B, C\}$. As we have already done with the standard General Lotto game, we propose a variation on the three-player General Lotto model that includes a preliminary stage (Stage 0) in which player B makes either a budgetary or battlefield concession. Below, we formalize the three stages of this variant, which we term the *three-player General Lotto game with concessions*, where it is assumed that the players’ actions in each stage become binding and common knowledge in subsequent stages:

- *Stage 0:* Player B selects one of the following concession formats:
 - *Budgetary concession:* Player B discards a portion of her budget $x \in (0, X_B]$; or,
 - *Battlefield concession:* Player B commits to allocating zero budget to a battlefield $b \in \mathcal{B}_B$.
- *Stage 1:* Player A allocates $X_{A,B}, X_{A,C} \geq 0$ of her budget to the fronts \mathcal{B}_B and \mathcal{B}_C , respectively, such that $X_{A,B} + X_{A,C} \leq X_A$ holds.

- *Stage 2:* Player A engages players B and C in the two resulting General Lotto games. If B made a budgetary concession of $x \in (0, X_B]$ in Stage 0, then A and B play the game $GL(X_{A,B}, X_B - x, \mathbf{v}_B)$. Else, if B made a battlefield concession of $b \in \mathcal{B}_B$, then A and B play the game $GL(X_{A,B}, X_B, \mathbf{v}_{B,-b})$, where $\mathbf{v}_{B,-b}$ denotes the vector of valuations for battlefields $\mathcal{B}_B \setminus \{b\}$. Players A and C play the game $GL(X_{A,C}, X_C, \mathbf{v}_C)$. Player A 's payoff is the sum of her expected payoffs in the two General Lotto games, and of v_b only if player B selected to concede the battlefield b in Stage 0. Each player $i \in \{B, C\}$ receives the expected payoff from her corresponding General Lotto game against A .

In order to identify player B 's optimal strategy in Stage 0 of the game we must first understand player A 's strategic behaviour in Stage 1. The allocation rule that maximizes A 's cumulative payoff in Stage 2 was characterized by Kovenock and Roberson [16]. We summarize this result below:

Fact 2. *Consider Stage 1 of the three-player General Lotto game where the players' budgets are normalized (w.l.o.g.) such that $X_A = 1$ and $X_B, X_C > 0$. Let $\Phi_B, \Phi_C > 0$ denote the cumulative value of non-conceded battlefields in the fronts \mathcal{B}_B and \mathcal{B}_C , respectively. Define $\mathcal{R}_{1i}, \mathcal{R}_{2i}, \mathcal{R}_{3i}$ and $\mathcal{R}_4, i \in \{B, C\}$ as the following regions:*

$$\begin{aligned} \mathcal{R}_{1i}(\Phi_i, \Phi_{-i}) &:= \{(X_i, X_{-i}) \text{ s.t. } \Phi_i/\Phi_{-i} > \max\{(X_i)^2, 1\}/(X_i X_{-i})\} \\ &\quad \cup \{(X_i, X_{-i}) \text{ s.t. } X_i < 1 \text{ and } \Phi_i/\Phi_{-i} = 1/(X_i X_{-i})\} \\ \mathcal{R}_{2i}(\Phi_i, \Phi_{-i}) &:= \{(X_i, X_{-i}) \text{ s.t. } \Phi_i/\Phi_{-i} > X_i/X_{-i} \text{ and } 0 < 1 - \sqrt{\Phi_i X_i X_{-i}/\Phi_{-i}} \leq X_{-i}\} \\ \mathcal{R}_{3i}(\Phi_i, \Phi_{-i}) &:= \{(X_i, X_{-i}) \text{ s.t. } \Phi_i/\Phi_{-i} \geq X_i/X_{-i} \text{ and } 1 - \sqrt{\Phi_i X_i X_{-i}/\Phi_{-i}} > X_{-i}\} \\ \mathcal{R}_4(\Phi_i, \Phi_{-i}) &:= \{(X_i, X_{-i}) \text{ s.t. } \Phi_i/\Phi_{-i} = X_i/X_{-i} \text{ and } X_i + X_{-i} \geq 1\}. \end{aligned}$$

Player A 's optimal allocation $X_{A,i}$ is determined in closed-form as follows:

- If $(X_i, X_{-i}) \in \mathcal{R}_{1i}(\Phi_i, \Phi_{-i})$, then $X_{A,i} = 1$.
- If $(X_i, X_{-i}) \in \mathcal{R}_{2i}(\Phi_i, \Phi_{-i})$, then $X_{A,i} = \sqrt{\Phi_i X_i X_{-i}/\Phi_{-i}}$.
- If $(X_i, X_{-i}) \in \mathcal{R}_{3i}(\Phi_i, \Phi_{-i})$, then $X_{A,i} = \sqrt{\Phi_i X_i}/(\sqrt{\Phi_i X_i} + \sqrt{\Phi_{-i} X_{-i}})$.
- If $(X_i, X_{-i}) \in \mathcal{R}_4(\Phi_i, \Phi_{-i})$, then any $X_{A,i} \in [1 - X_{-i}, X_i]$ is optimal,

where $X_{A,-i} = 1 - X_{A,i}$ in all the above cases.

Observe that the result above can be applied in Stage 1, whether or not player B makes a concession. If player B makes no concessions (i.e., the nominal game), then her payoff in Stage 2 is $\Phi_B \cdot L(X_B, X_{A,B})$, where we use $X_{A,B}$ here to denote player A 's optimal allocation to the front \mathcal{B}_B in Stage 1 when there is no concession. Otherwise, if player B makes a budgetary concession of $x \in (0, X_B]$, then her payoff in Stage 2 is $\Phi_B \cdot L(X_B - x, X'_{A,B})$, and if player B makes a battlefield concession of $b \in \mathcal{B}_B$, then her payoff in Stage 2 is $(\Phi_B - v_b) \cdot L(X_B, X''_{A,B})$, where we use $X'_{A,B}$ and $X''_{A,B}$ here to denote player A 's optimal allocation in Stage 1 to the General Lotto game against player B in response to the budgetary and battlefield concessions, respectively. Crucially, observe that player B translates

the point (X_B, X_C) to the left by making budgetary concessions, and alters the parametric regions identified in Fact 2 by making battlefield concessions.

The following observations will be important in the proofs of the forthcoming results:

- i. $X_{A,B} > X_B$ holds in regions \mathcal{R}_{1B} (if $X_B < 1$), \mathcal{R}_{2B} , \mathcal{R}_{3B} and \mathcal{R}_{3C} , while $X_{A,B} \leq X_B$ holds in regions \mathcal{R}_{1B} (if $X_B \geq 1$), \mathcal{R}_{1C} , \mathcal{R}_{2C} and \mathcal{R}_4 .
- ii. The closed-form expressions of player A 's optimal allocation – and, thus, all of the players' payoffs – are identical in regions \mathcal{R}_{3B} and \mathcal{R}_{3C} . Thus, it is equivalent to denote the union of the parametric regions $\mathcal{R}_{3i}(\Phi_i, \Phi_{-i})$, $i \in \{B, C\}$, simply by $\mathcal{R}_3(\Phi_B, \Phi_C)$. Further, note that any point (X_B, X_C) in \mathcal{R}_3 must satisfy $X_B + X_C < 1$ since $X_{A,i} > X_i$, $i \in \{B, C\}$.
- iii. As shown in Fig. 2(a), the point (X_B, X_C) translates to the left in the (X_B, X_C) -plane following a budgetary concession by player B . On the other hand, the regions identified in Fact 2 remain unperturbed.
- iv. In contrast, the concession of a battlefield with value \hat{v} does not translate the point (X_B, X_C) , but rather modifies the regions identified in Fact 2, as shown in Fig. 2(b). In particular, the line $X_C = \Phi_C X_B / \Phi_B$ which serves as the boundary between the regions $\mathcal{R}_{1B} \cup \mathcal{R}_{2B}$ and $\mathcal{R}_{1C} \cup \mathcal{R}_{2C}$ (i.e., the *median line*) rotates counterclockwise about the origin following a battlefield concession by player B . Crucially, since the points on the (X_B, X_C) -plane remain stationary, a point (X_B, X_C) can move from one region to another (i.e., *transit*) following a battlefield concession by either player.

3 Main Results

All of the results in this section focus on concessions in the three-player General Lotto game from the perspective of player B . However, by flipping the players' labels, all the results apply identically to concessions from the perspective of player C . Throughout this section, we refer to concessions that strictly improve the player's payoff above her payoff in the nominal setting as beneficial budgetary and battlefield concessions.

Budgetary Concessions. We first focus on budgetary concessions, and show that players cannot improve their payoffs by making such concessions.

Theorem 1. *Consider the three-player General Lotto game with $X_A = 1$, $X_B, X_C \geq 0$ and $\Phi_B, \Phi_C \geq 0$. Player B cannot benefit from a budgetary concession.*

We present the proof of Theorem 1 in Appendix A, for ease of presentation. As the proof is fairly technical, we provide an intuitive interpretation for the reader's convenience. Suppose player B makes a budgetary concession of $x \in (0, X_B]$. Observe that the budgetary concession leaves player B more vulnerable to attacks from player A , since her budget is lowered, but the cumulative value of the battlefields in front \mathcal{B}_B remains unchanged. As a result, the adversary

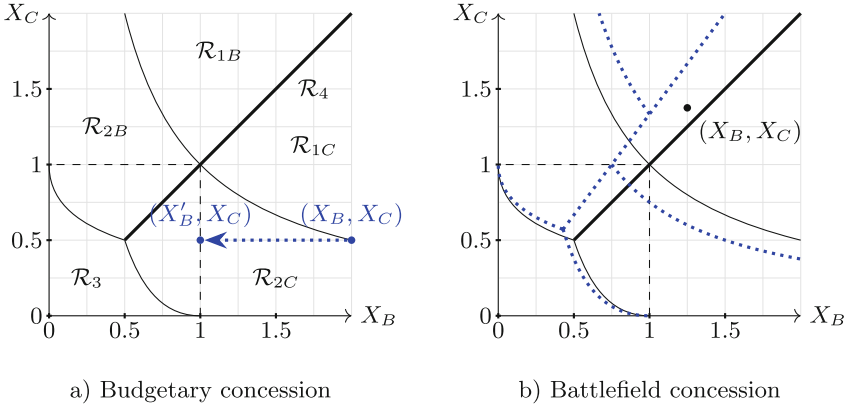


Fig. 2. The regions dividing the possible player budgets (X_B, X_C) in Stage 2, as derived in [16] and reviewed in Fact 2. (a) Illustration of the Stage 1 regions in the three-player General Lotto game with $\Phi_B = \Phi_C$. The solid, black lines depict the borders between the labelled regions. In blue, we depict the impact of a budgetary concession: the point $(X_B, X_C) = (2, 0.5)$ translates to the left to $(X_B - x, X_C) = (1, 0.5)$ after player B makes a budgetary concession of $x = 1$. (b) We depict the impact of a battlefield concession within the same setting as (a), but where B makes a battlefield concession of b with $v_b = \Phi_C/4$. The solid, black lines depict the borders between the regions for no concession (i.e., $\mathcal{R}_j(\Phi_B, \Phi_C)$), while the dotted, blue lines depict the borders between the regions after the concession of b (i.e., $\mathcal{R}_j(\Phi_B - v_b, \Phi_C)$). Observe that all points on the plot, including $(X_B, X_C) = (1.25, 1.375)$, remain stationary, while the regions change. Notably, (X_B, X_C) is in region \mathcal{R}_{1B} if no concession is made, and in \mathcal{R}_{1C} after the concession of battlefield b .

will seek either the same or greater payoff from the front \mathcal{B}_B . In the best case, the amount of payoff that the adversary extracts from the front \mathcal{B}_B will stay the same, as is the case if the pairs (X_B, X_C) and $(X_B - x, X_C)$ are both in $\mathcal{R}_{1C}(\Phi_B, \Phi_C)$, i.e., player A still sends no budget to \mathcal{B}_B . In all other settings, player B 's payoff will strictly decrease after a budgetary concession.

Battlefield Concessions. Next, we focus on battlefield concessions. Here, we are concerned with identifying instances in which a battlefield concession is beneficial for player B , i.e., B 's resulting payoff is higher than in the nominal game. In particular, we seek conditions on the budgets X_A, X_B , and X_C , and the players' front values Φ_B and Φ_C for which there exists a beneficial battlefield concession. Note here that we are not concerned with the particular vectors of battlefield valuations $\mathbf{v}_B, \mathbf{v}_C$ that constitute each front. As such, we allow player B to have full choice over the conceded value $v \in [0, \Phi_B]$. Our next result identifies sufficient conditions for the existence of beneficial battlefield concessions in any three-player General Lotto game.

Theorem 2. Consider the three-player General Lotto game with $X_A = 1$, $X_B, X_C \geq 0$ and $\Phi_B, \Phi_C \geq 0$. Let $v^* = \Phi_B - \Phi_C X_B / X_C$. The following conditions characterize sufficient conditions under which player B has a beneficial battlefield concession of value $v_b = v^*$:

- (i) If $(X_B, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$ and $X_B, X_C \geq 1$, then $v^* < \Phi_B / (2X_B)$;
- (ii) If $(X_B, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$, $X_B \geq 1$ and $X_C < 1$, then

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B} \right] > \Phi_B \cdot \left(1 - \frac{1}{2X_B} \right);$$

- (iii) If $(X_B, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$, $X_B < 1$ and $X_C \geq 1$, then $v^* < \Phi_B \cdot (1 - X_B / 2)$;
- (iv) If $(X_B, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$, and $X_B, X_C < 1$, then

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B} \right] > \frac{\Phi_B X_B}{2}.$$

- (v) If $(X_B, X_C) \in \mathcal{R}_{2B}(\Phi_B, \Phi_C)$ and $X_C \geq 1$, then

$$v^* < \Phi_B \cdot \left[1 - \frac{X_B}{2\sqrt{\Phi_B X_B X_C / \Phi_C}} \right]; \text{ and,}$$

- (vi) If $(X_B, X_C) \in \mathcal{R}_{2B}(\Phi_B, \Phi_C)$, $X_C < 1$ and $X_B + X_C \geq 1$, then

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B} \right] > \frac{\Phi_B X_B}{2\sqrt{\Phi_B X_B X_C / \Phi_C}}.$$

We present the proof of Theorem 2 in Appendix B, for ease of presentation. In place of the proof, we devote the remainder of this section to developing the intuition about this positive result.

First, we explain the significance of the value v^* defined in the claim of Theorem 2. Observe that v^* is precisely the battlefield value that satisfies $(\Phi_B - v^*) / X_B = \Phi_C / X_C$. Thus, when $X_B + X_C \geq 1$ and the point (X_B, X_C) is nominally in one of the regions $\mathcal{R}_{1B}(\Phi_B, \Phi_C)$ or $\mathcal{R}_{2B}(\Phi_B, \Phi_C)$, player B can concede a battlefield of value $v_b = v^* + \epsilon$, $\epsilon \rightarrow 0^+$, to alter the regions in such a way that (X_B, X_C) is in either $\mathcal{R}_{1C}(\Phi_B - v_b, \Phi_C)$ (when $X_C \geq 1$) or $\mathcal{R}_{2C}(\Phi_B - v_b, \Phi_C)$ (when $X_C < 1$). Note that if $X_B + X_C < 1$ instead, then the concession of the battlefield of value v_b satisfies $(X_B, X_C) \in \mathcal{R}_3(\Phi_B - v_b, \Phi_C)$, and does not offer any benefit to B.

Next, we consider simulation results identifying the parameter regime in which our conditions for beneficial battlefield concessions hold. In Fig. 3, we plot player B's optimal battlefield concession, where the players' budgets are normalized such that $X_A = 1$, and $X_B, X_C \in [0, 1.2]$. In each of the panels, the cumulative values of battlefields in the two fronts are as follows: Fig. 3a has $\Phi_B = 1, \Phi_C = 0.5$, Fig. 3b has $\Phi_B = 1, \Phi_C = 1$, and Fig. 3c has $\Phi_B = 1, \Phi_C = 2$.

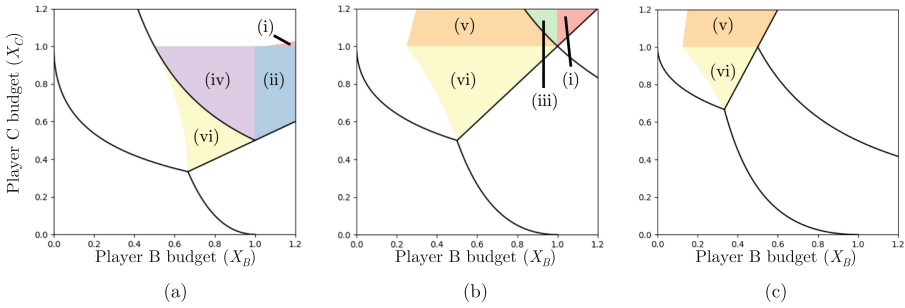


Fig. 3. Existence of beneficial battlefield concessions for player B . The coloured regions correspond with our sufficient conditions for beneficial battlefield concessions for player B under normalized player budgets (i.e., $X_A = 1$), for $X_B, X_C \in [0, 1.2]$ and (a) $\Phi_B = 1, \Phi_C = 0.5$, (b) $\Phi_B = 1, \Phi_C = 1$, and (c) $\Phi_B = 1, \Phi_C = 2$. The white area coincide with the points (X_B, X_C) for which our sufficient conditions are not met. The solid, black lines divide the (X_B, X_C) -plane into the various regions, where \mathcal{R}_{1C} , \mathcal{R}_{2C} and \mathcal{R}_3 are as labelled in plot (a), \mathcal{R}_{2B} is at the top left of each plot (not labelled), and \mathcal{R}_{1B} does not appear. In each of the coloured areas, it is beneficial for player B to concede a battlefield of value v^* , and the different coloured areas' labels coincide with Conditions (i)–(vi), all of which are defined in the claim of Theorem 2. (Color figure online)

Player B has no beneficial concession in the white area. The various regions \mathcal{R}_j , defined in Fact 2, are divided by solid, black lines. The coloured areas' labels coincide with the conditions identified in the claim of Theorem 2.

As seen in Fig. 3 and the definitions of Conditions (i)–(vi), the set of beneficial battlefield concessions we identify appear in regions where the ratio between the cumulative value of the battlefields in front \mathcal{B}_B and B 's initial budget endowment, Φ_B/X_B , is greater than the ratio Φ_C/X_C , and the players B and C together possess enough budget to force A to prioritize one of her General Lotto games over the other (i.e., $\Phi_B/X_B > \Phi_C/X_C$ and $X_B + X_C \geq 1$). In such scenarios, player A primarily pursues her General Lotto game against B in the nominal game. By conceding enough battlefield value $v \in [0, \Phi_B]$ such that $(\Phi_B - v)/X_B < \Phi_C/X_C$, player B can force A to prioritize her game against C instead. Interestingly, if the difference between Φ_B/X_B and Φ_C/X_C is moderate, then the gains from shifting A 's attention outweigh the loss of the forfeited battlefield's value. If the difference between Φ_B/X_B and Φ_C/X_C is too high, however, then too much value v must be conceded by B to minimize her conflict with A , and the gains will not outweigh the losses.

In Sect. 1.3, we briefly describe the variant of the three-player General Lotto game studied in [16]. Recall that, in their variant, the players B and C have the opportunity to negotiate an alliance which entails a unilateral transfer of budgetary resources in Stage 0 of the game, and that cases are identified in which forming an alliance is mutually beneficial for B and C . The results in [16] suggest that mutually beneficial alliances only occur when the difference

between the ratios Φ_B/X_B and Φ_C/X_C is sufficiently large. In contrast, our findings show that beneficial battlefield concessions only exist when the ratios Φ_B/X_B and Φ_C/X_C are close. This comparison suggests that, if there are significant asymmetries in the players' strengths relative to the values of their respective contests, then cooperative mechanisms, such as alliances, provide strategic advantages; meanwhile, if differences in players' relative strengths are small, then unilateral mechanisms such as battlefield concessions prevail.

4 Conclusions and Future Work

In this paper, we considered the viability of "concessions" as a component of strategic decision-making in adversarial environments. We considered two types of concessions: *budgetary concessions*, where a competitor voluntarily reduces one's resource budget, and *battlefield concessions*, where a player voluntarily forfeits a certain prize to her adversary. Intuitively, concessions should not offer strategic advantages as they weaken one's competitive position. However, we demonstrated that they do offer benefits if made correctly. We studied concessions under the framework of General Lotto games, where we showed that neither type of concession offers benefits under the two-player model. However, in settings where two players compete against a common adversary, we showed that one of the two players can often improve her payoff by conceding a battlefield to the adversary.

This work provides several avenues for future study. First, we have shown that conceding battlefields is beneficial when General Lotto games are the underlying model of conflict. However, we suspect this phenomenon is robust to larger classes of models, e.g., Tullock and other contest success functions. Second, considering a richer setting wherein both players can simultaneously make concessions to the adversary opens questions of what strategic outcomes are possible. Finally, though we have studied concessions as a strategic component in two- and three-player settings, broader forms of strategic pre-commitments and more general interaction networks could be considered.

A Proof of Theorem 1

Proof. The proof amounts to showing that player B 's payoff is nonincreasing for any budgetary concession $x \in (0, X_B]$ such that $(X_B - x, X_C)$ is in any of the regions \mathcal{R}_j .

We first consider the scenario where $(X_B, X_C) \in \mathcal{R}_{1C}(\Phi_B, \Phi_C)$. Recall that, in this scenario, player A commits no budget to the battlefields in the front \mathcal{B}_B . Thus, player B 's payoff before the concession is Φ_B , the highest possible payoff. Furthermore, $(X_B - x, X_C) \in \mathcal{R}_{1C}(\Phi_B, \Phi_C)$ can only hold if $(X_B, X_C) \in \mathcal{R}_{1C}(\Phi_B, \Phi_C)$ as well, since the value $1 - \sqrt{\Phi_C(X_B - x)X_C/\Phi_B}$ is increasing in x . If $(X_B - x, X_C) \in \mathcal{R}_4(\Phi_B, \Phi_C)$, then any budgetary concession $x' < x$ would be in either \mathcal{R}_{1C} or \mathcal{R}_{2C} , since $X_B + X_C \geq 1$ in \mathcal{R}_4 , while any budgetary concession $x' > x$ would be in either \mathcal{R}_{1B} , \mathcal{R}_{2B} or \mathcal{R}_3 . Thus, conceding any amount $x' < x$

would guarantee B greater payoff since $X_{A,B} = 1 - X_{A,C}$, and $X_{A,C} = 1$ in \mathcal{R}_{1C} and $X_{A,C} > X_C$ in \mathcal{R}_{2C} whereas $X_{A,C} \in [\max\{0, 1 - X_B\}, \min\{1, X_C\}]$ in \mathcal{R}_4 . Further, conceding any amount $x' > x$ cannot guarantee B greater payoff since $X_{A,B} = 1$ in \mathcal{R}_{1B} , and $X_{A,B} > X_B$ in \mathcal{R}_{2B} and \mathcal{R}_3 , whereas $X_{A,B} \in [\max\{0, 1 - X_C\}, \min\{1, X_B\}]$ in \mathcal{R}_4 . In all other regions, we show that player B 's payoff is strictly decreasing in x by checking the partial derivative with respect to $x \geq 0$:

If $(X_B - x, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$ and $X_B - x > 1$, then

$$\frac{\partial}{\partial x} \Phi_B \left[1 - \frac{1}{2(X_B - x)} \right] = -\frac{\Phi_B}{2(X_B - x)^2} < 0.$$

Else, if $(X_B - x, X_C) \in \mathcal{R}_{1B}(\Phi_B, \Phi_C)$ and $X_B - x \leq 1$, then

$$\frac{\partial}{\partial x} \frac{\Phi_B(X_B - x)}{2} = -\frac{\Phi_B}{2} < 0.$$

If $(X_B - x, X_C) \in \mathcal{R}_{2B}(\Phi_B, \Phi_C)$, then

$$\frac{\partial}{\partial x} \frac{\Phi_B(X_B - x)}{2\sqrt{\frac{\Phi_B(X_B - x)X_C}{\Phi_C}}} = -\frac{\Phi_B}{4\sqrt{\frac{\Phi_B(X_B - x)X_C}{\Phi_C}}} < 0.$$

If $(X_B - x, X_C) \in \mathcal{R}_{2C}(\Phi_B, \Phi_C)$, then

$$\frac{\partial}{\partial x} \Phi_B \left[1 - \frac{1 - \sqrt{\frac{\Phi_C(X_B - x)X_C}{\Phi_B}}}{2(X_B - x)} \right] = -\frac{\Phi_B \left[2 - \sqrt{\frac{\Phi_C(X_B - x)X_C}{\Phi_B}} \right]}{4(X_B - x)^2} < 0,$$

which is strictly negative as the condition $1 - \sqrt{\Phi_C(X_B - x)X_C/\Phi_B} \geq 0$ must hold in \mathcal{R}_{2C} . Finally, if $(X_B - x, X_C) \in \mathcal{R}_3(\Phi_B, \Phi_C)$, then

$$\frac{\partial}{\partial x} \frac{\Phi_B(X_B - x)}{2\frac{\sqrt{\Phi_B(X_B - x)}}{\sqrt{\Phi_B(X_B - x)} + \sqrt{\Phi_C X_C}}} = -\frac{\Phi_B}{2} - \frac{\Phi_B \sqrt{\Phi_C X_C}}{4\sqrt{\Phi_B(X_B - x)}} < 0.$$

This concludes the proof. □

B Proof of Theorem 2

Before presenting the proof, we note that, in the case of battlefield concessions, we can disregard the scenario when $(X_B, X_C) \in \mathcal{R}_4(\Phi_B - v, \Phi_C)$, for any $v \in [0, \Phi_B]$. To see why, consider a battlefield concession of value v such that (X_B, X_C) is in \mathcal{R}_4 , i.e., $X_B + X_C \geq 1$ and $(\Phi_B - v)/X_B = \Phi_C/X_C$. Observe that by conceding a battlefield of value slightly greater than v (i.e., $v' = v + \epsilon$ for $\epsilon \rightarrow 0^+$), player B obtains strictly higher payoff as (X_B, X_C) now falls in region \mathcal{R}_{1C} (if $X_C \geq 1$) or region \mathcal{R}_{2C} (if $X_C < 1$). Thus, in the following

proof, we assume that any point (X_B, X_C) with $X_B + X_C \geq 1$ will transit directly from $\mathcal{R}_{1B}(\Phi_B - v, \Phi_C)$ (if $X_B \geq 1$) or $\mathcal{R}_{2B}(\Phi_B - v, \Phi_C)$ (if $X_B < 1$), to $\mathcal{R}_{1C}(\Phi_B - v, \Phi_C)$ (if $X_C \geq 1$) or $\mathcal{R}_{2C}(\Phi_B - v, \Phi_C)$ (if $X_C < 1$), as v is increased, without first passing through \mathcal{R}_4 .

Proof. The proof amounts to verifying that the conditions laid out in the claim guarantee that player B 's payoff after the battlefield concession is greater than her payoff in the nominal three-player General Lotto game. Before we continue, it is critical to note that v^* as defined in the claim is precisely the value that satisfies $(\Phi_B - v^*)/X_B = \Phi_C/X_C$. Thus, for $(X_B, X_C) \in \mathcal{R}_{1B} \cup \mathcal{R}_{2B}$ and $X_B + X_C \geq 1$, the battlefield concession of value v^* satisfies $(X_B, X_C) \in \mathcal{R}_{1C} \cup \mathcal{R}_{2C}$. We present the remainder of the proof in parts corresponding with each of the conditions in the claim.

Conditions (i): The point (X_B, X_C) is nominally in the region $\mathcal{R}_{1B}(\Phi_B, \Phi_C)$ with $X_B \geq 1$, and, thus, player B 's nominal payoff is $\Phi_B \cdot (1 - 1/2X_B)$. Since $X_C \geq 1$, the battlefield concession of value v^* satisfies $(X_B, X_C) \in \mathcal{R}_{1C}(\Phi_B - v^*, \Phi_C)$, and player B 's resulting payoff is $\Phi_B - v^*$. It follows that the battlefield concession of value v^* benefits player B if

$$\Phi_B - v^* > \Phi_B \cdot \left(1 - \frac{1}{2X_B}\right).$$

Rearranging the above inequality gives the condition in the claim.

Condition (ii): Once again, player B 's nominal payoff is $\Phi_B \cdot (1 - 1/2X_B)$. Since $X_C < 1$ and $X_B + X_C \geq 1$, the battlefield concession of value v^* satisfies $(X_B, X_C) \in \mathcal{R}_{2C}(\Phi_B - v^*, \Phi_C)$, and player B 's resulting payoff is $(\Phi_B - v^*) \cdot [1 - (1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}) / (2X_B)]$. It follows that the battlefield concession of value v^* benefits player B if

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B}\right] > \Phi_B \cdot \left(1 - \frac{1}{2X_B}\right).$$

Condition (iii): Observe that this condition resembles Condition (i), except that $X_B < 1$. Thus, the only difference is that player B 's nominal payoff is $\Phi_B X_B / 2$. It follows that the battlefield concession of value v^* benefits player B if

$$\Phi_B - v^* > \frac{\Phi_B X_B}{2}.$$

Rearranging the above inequality gives the condition in the claim.

Condition (iv): Observe that this condition resembles Condition (iii), except that $X_C < 1$. Thus, the only difference is that player B 's resulting payoff is $(\Phi_B - v^*) \cdot [1 - (1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}) / (2X_B)]$, as in Condition (ii). It follows that the battlefield concession of value v^* benefits player B if

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B}\right] > \frac{\Phi_B X_B}{2\sqrt{\Phi_B X_B X_C / \Phi_C}}.$$

Condition (v): The point (X_B, X_C) is nominally in the region $\mathcal{R}_{2B}(\Phi_B, \Phi_C)$, and, thus, player B 's nominal payoff is $\Phi_B \cdot X_B / (2\sqrt{\Phi_B X_B X_C / \Phi_C})$ since $X_B < 1$ must hold in \mathcal{R}_{2B} . Since $X_C \geq 1$, player B 's resulting payoff after the battlefield concession of value v^* is $\Phi_B - v^*$, as in Condition (i). It follows that the battlefield concession of value v^* benefits player B if

$$\Phi_B - v^* > \frac{\Phi_B X_B}{2\sqrt{\Phi_B X_B X_C / \Phi_C}}.$$

Rearranging the above inequality gives the condition in the claim.

Condition (vi): This condition resembles Condition (iv), except that $X_C < 1$. Thus, the only difference is that player B 's resulting payoff is $(\Phi_B - v^*) \cdot [1 - (1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}) / (2X_B)]$, as in Condition (ii). It follows that the battlefield concession of value v^* benefits player B if

$$(\Phi_B - v^*) \cdot \left[1 - \frac{1 - \sqrt{\Phi_C X_B X_C / (\Phi_B - v^*)}}{2X_B} \right] > \frac{\Phi_B X_B}{2\sqrt{\Phi_B X_B X_C / \Phi_C}}.$$

This concludes the proof. \square

References

1. Bell, R.M., Cover, T.M.: Competitive optimality of logarithmic investment. *Math. Oper. Res.* **5**(2), 161–166 (1980)
2. Boix-Adserà, E., Edelman, B.L., Jayanti, S.: The multiplayer colonel blotto game. In: *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 47–48 (2020)
3. Borel, E.: La théorie du jeu les équations intégrales à noyau symétrique. *C. R. Acad.* **173** (1921)
4. Chandan, R., Paarporn, K., Marden, J.R.: When showing your hand pays off: announcing strategic intentions in colonel blotto games. In: *2020 American Control Conference (ACC)*, pp. 4632–4637 (2020)
5. Ferdowsi, A., Saad, W., Mandayam, N.B.: Colonel blotto game for sensor protection in interdependent critical infrastructure. *IEEE Internet Things J.* **8**(4), 2857–2874 (2021)
6. Friedman, L.: Game-theory models in the allocation of advertising expenditures. *Oper. Res.* **6**(5), 699–709 (1958)
7. Fuchs, Z.E., Khargonekar, P.P.: A sequential colonel blotto game with a sensor network. In: *2012 American Control Conference (ACC)*, pp. 1851–1857. IEEE (2012)
8. Goyal, S., Vigier, A.: Attack, defence, and contagion in networks. *Rev. Econ. Stud.* **81**(4), 1518–1542 (2014)
9. Gross, O., Wagner, R.: A continuous Colonel Blotto game. Technical report, RAND Project, Air Force, Santa Monica (1950)
10. Guan, S., Wang, J., Yao, H., Jiang, C., Han, Z., Ren, Y.: Colonel Blotto games in network systems: models, strategies, and applications. *IEEE Trans. Netw. Sci. Eng.* **7**(2), 637–649 (2020)

11. Gupta, A., Başar, T., Schwartz, G.A.: A three-stage colonel blotto game: when to provide more information to an adversary. In: Poovendran, R., Saad, W. (eds.) *GameSec 2014*. LNCS, vol. 8840, pp. 216–233. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12601-2_12
12. Gupta, A., Schwartz, G., Langbort, C., Sastry, S.S., Basar, T.: A three-stage colonel blotto game with applications to cyberphysical security. In: 2014 American Control Conference, pp. 3820–3825 (2014)
13. Hart, S.: Discrete colonel blotto and general lotto games. *Int. J. Game Theory* **36**(3–4), 441–460 (2008)
14. Kovenock, D., Roberson, B.: Conflicts with multiple battlefields. In: Garfinkel, M., Skaperdas, S. (eds.) *The Oxford Handbook of the Economics of Peace and Conflict*. Oxford University Press, Oxford (2012)
15. Kovenock, D., Roberson, B.: Generalizations of the general lotto and colonel blotto games. *Econ. Theory* **71**(3), 997–1032 (2020). <https://doi.org/10.1007/s00199-020-01272-2>
16. Kovenock, D., Roberson, B.: Coalitional colonel blotto games with application to the economics of alliances. *J. Pub. Econ. Theory* **14**(4), 653–676 (2012)
17. Macdonell, S.T., Mastronardi, N.: Waging simple wars: a complete characterization of two-battlefield blotto equilibria. *Econ. Theory* **58**(1), 183–216 (2015)
18. Mattioli, D., Lombardo, C.: Amazon met with startups about investing, then launched competing products. *Wall Street J.* (2020)
19. Myerson, R.B.: Incentives to cultivate favored minorities under alternative electoral systems. *Am. Polit. Sci. Rev.* **87**(4), 856–869 (1993)
20. Pita, J., et al.: Deployed armor protection: the application of a game theoretic model for security at the Los Angeles international airport. In: *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems: Industrial Track*, pp. 125–132 (2008)
21. Powell, R.: Sequential, nonzero-sum “blotto”: allocating defensive resources prior to attack. *Games Econ. Behav.* **67**(2), 611–615 (2009)
22. Rinott, Y., Scarsini, M., Yu, Y.: A colonel blotto gladiator game. *Math. Oper. Res.* **37**(4), 574–590 (2012)
23. Roberson, B.: The colonel blotto game. *Econ. Theory* **29**(1), 1–24 (2006)
24. Schwartz, G., Loiseau, P., Sastry, S.S.: The heterogeneous Colonel Blotto game. In: *International Conference on NETWORK Games, CONTROL and OPTimization*, pp. 232–238 (2014)
25. Shahrivar, E.M., Sundaram, S.: Multi-layer network formation via a colonel blotto game. In: 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 838–841 (2014)
26. Snyder, J.M.: Election goals and the allocation of campaign resources. *Econometrica: J. Econ. Soc.*, 637–660 (1989)
27. Tambe, M.: *Security and Game Theory: Algorithms, Deployed Systems*. Cambridge University Press, Cambridge (2011). *Lessons Learned*
28. Thomas, C.: N-dimensional blotto game with heterogeneous battlefield values. *Econ. Theory* **65**(3), 509–544 (2018)
29. Vu, D.Q., Loiseau, P.: Colonel blotto games with favoritism: Competitions with pre-allocations and asymmetric effectiveness. arXiv preprint [arXiv:2106.00617](https://arxiv.org/abs/2106.00617) (2021)
30. Yang, R., Ford, B., Tambe, M., Lemieux, A.: Adaptive resource allocation for wildlife protection against illegal poachers. In: *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 453–460 (2014)



Providing Slowdown Information to Improve Selfish Routing

Philip N. Brown^(✉) 

University of Colorado Colorado Springs, Colorado Springs, CO, USA
philip.brown@uccs.edu

Abstract. Recent research in the social sciences has identified situations in which small changes in the way that information is provided to consumers can have large aggregate effects on behavior. This has been promoted in popular media in areas of public health and wellness, but its application to other areas has not been broadly studied. This paper presents a simple model which expresses the effect of providing commuters with carefully-curated information regarding aggregate traffic “slowdowns” on the various roads in a transportation network. Much of the work on providing information to commuters focuses specifically on travel-time information. However, the model in the present paper allows a system planner to provide slowdown information as well; that is, commuters are additionally told how much slower each route is as compared to its uncongested state. We show that providing this additional information can improve equilibrium routing efficiency when compared to the case when commuters are only given information about travel time, but that these improvements in congestion are not universal. That is, transportation networks exist on which any provision of slowdown information can harm equilibrium congestion. In addition, this paper illuminates a deep connection between the effects of commuter slowdown-sensitivity and the study of marginal-cost pricing and altruism in congestion games.

Keywords: Congestion Game · Traffic Information System · Transportation Networks

1 Introduction

Today, in the age of the internet of things, big data, and pervasive computing, engineered systems are becoming increasingly interconnected with the human populations that they serve. System-level performance is directly affected by the choices of human users, customers, and adversaries, and it has long been recognized that self-interested behavior by users may lead to gross system inefficiencies [21, 22]. Thus, both the importance and the feasibility of influencing human behavior in intelligent ways are increasing simultaneously.

This work is supported in part by NSF Grant ECCS-2013779.

A popular modeling framework to study methods of influencing selfish behavior in engineered systems is the *nonatomic congestion game*, which is often used to model urban transportation networks [2]. This model was one of the first to admit straightforward characterizations of the *price of anarchy*, a popular measure of the social cost of selfish behavior [22]. It is also a natural setting to study various modified models of human decision-making; examples include altruism [10], pessimism [19], risk-aversion [17], and various generalizations [15]. More relevant to this paper, it has proved fertile ground for the study of various methods of influencing user behavior: taxation [4, 7, 12, 13], autonomously-controlled traffic [3, 16], and traffic information systems [1, 11, 18] have all been investigated in this context. The majority of this literature depends on the usual game-theoretic assumption that users are expected utility maximizers.

In parallel, an increasingly popular conceptual framework for studying the practice of influencing human behavior is that of *behavioral economics*. This framework grew out of empirical studies which showed that humans tend to deviate from the behaviors prescribed by expected utility maximization, and that these deviations manifest themselves in predictable ways [25]. If humans are predictably irrational, it seems reasonable that engineers might attempt to exploit this predictability to influence behavior. This idea has led to the concept of the “nudge”—the popular name given to the idea that making slight modifications to a person’s environment can have a profound effect on their behavior [24].

This paper draws inspiration from the concept of the nudge, and asks if drivers in transportation networks could perhaps be influenced by providing simple pieces of information that are carefully designed to exploit behavioral biases in their decision-making process. As a simple preliminary model for studying such questions, this paper begins with the typical assumption that drivers prefer low travel times over high travel times. Some drivers, however, additionally dislike the idea of choosing a route that exhibits higher travel time than its uncongested free-flow travel time. That is, some drivers are slowdown-averse. A system planner, knowing about the existence (but perhaps not the magnitude) of this slowdown-aversion, may wish to exploit it to improve aggregate congestion, and provides a signal informing drivers of the presence and magnitude of slowdowns on various routes. Note that unlike much other work on traffic information systems, we do not require ordinary drivers to perform any type of Bayesian inference; rather, our model captures a phenomenon in which each driver acts in response to her “gut feel” about how much she dislikes traffic slowdowns.

The main question of this paper is this: when can a planner be certain that providing slowdown information to drivers will improve congestion? In Theorem 1, we show that networks exist for which any slowdown information provided to drivers increases equilibrium delays, and thus on these networks, the planner’s only reasonable course of action is to provide no slowdown signal at all. Nonetheless, Theorem 1 also shows that on parallel networks, slowdown information can always reduce equilibrium delays. Perhaps unintuitively, these results

are tightly connected with the literature on marginal-cost pricing and altruistic behavior in transportation networks [5, 10]. Indeed, this condition can be stated as a sufficient condition for improvements due to slowdown signaling: If altruism increases equilibrium delays on some network, then slowdown signaling also increases equilibrium delays on that network.

Following this, Theorem 2 essentially asks the question “is it optimal for planners to provide accurate slowdown information to drivers?” Alternatively, what is the optimal slowdown signal with respect to minimizing equilibrium congestion? Here, we consider an example setting in which the planner has significant levels of information about the population’s slowdown preferences, as well as some minimal information about aggregate demand. In this setting, Theorem 2 shows that planners with access to this additional information optimally *overstate* the severity of network slowdowns. That is, a planner’s optimal course of action may well be to deliberately provide misleading information to drivers.

Lastly, the paper closes with a note on the implications that this work has for the broader study of heterogeneous nonatomic congestion games. Here, it is shown that many of the well-studied forms of heterogeneous congestion games appearing in the literature belong to a particularly well-behaved class of games known as *weighted potential games*. That is, the Nash equilibria of these games can be characterized as the maximizers of a single global convex potential function. From an analytical point of view, the ramifications of this are broad: the existing work on these types of games depends on *ad hoc* characterizations of equilibrium behavior, and the existence of a potential function may unify these approaches.

2 Model and Performance Metrics

2.1 Routing Game

Consider a network routing problem for a network (V, E) comprised of vertex set V and edge set E . We call a source/destination vertex pair $(\sigma^c, t^c) \in (V \times V)$ a *commodity*, and the set of all such commodities \mathcal{C} . For each commodity $c \in \mathcal{C}$, there is a mass of traffic $r^c > 0$ that needs to be routed from σ^c to t^c . We write $\mathcal{P}^c \subset 2^E$ to denote the set of *paths* available to traffic in commodity c , where each path $p \in \mathcal{P}^c$ consists of a set of edges connecting σ^c to t^c . Let $\mathcal{P} = \cup \{\mathcal{P}^c\}$. A network is called *symmetric* if there is exactly one commodity: $\mathcal{C} = \{c\}$, i.e., all traffic routes from a common source σ to a common destination t using a common path set \mathcal{P} . A network is called a *parallel* network if all commodities share a single source-destination pair and all paths are disjoint; i.e., for all paths $p, p' \in \mathcal{P}$, $p \cap p' = \emptyset$. Note that a parallel network need not be symmetric; although all traffic must share a common source and destination, the path sets \mathcal{P}^c available to traffic from different commodities may differ.

We write $f_p^c \geq 0$ to denote the mass of traffic from commodity c using path p , and $f_p := \sum_{c \in \mathcal{C}} f_p^c$. A *feasible flow* $f \in \mathbb{R}^{|\mathcal{P}|}$ is an assignment of traffic to various paths such that for each c , $\sum_{p \in \mathcal{P}^c} f_p^c = r^c$ and $\sum_{c \in \mathcal{C}} r^c = r$.

Given a flow f , the flow on edge e is given by $f_e = \sum_{p:e \in p} f_p$. To characterize transit delay as a function of traffic flow, each edge $e \in E$ is associated with a specific latency function $\ell_e : [0, r] \rightarrow [0, \infty)$; $\ell_e(f_e)$ denotes the delay experienced by users of edge e when the edge flow is f_e . We adopt the standard assumptions that each latency function is nondecreasing, convex, and continuously differentiable. We measure the cost of a flow f by the *total latency*, given by

$$\mathcal{L}(f) = \sum_{e \in E} f_e \cdot \ell_e(f_e) = \sum_{p \in \mathcal{P}} f_p \cdot \ell_p(f), \quad (1)$$

where $\ell_p(f) = \sum_{e \in p} \ell_e(f_e)$ denotes the latency on path p . We denote the flow that minimizes the total latency by

$$f^* \in \underset{f \text{ is feasible}}{\operatorname{argmin}} \mathcal{L}(f). \quad (2)$$

A *routing problem* is given by $G = (V, E, \mathcal{C}, \{\ell_e\})$. Classes of routing problems are denoted by \mathcal{G} . For $d \geq 1$, we write \mathcal{G}^d to denote the class of all routing problems with latency functions of the form $\ell_e(f_e) = a_e(f_e)^d + b_e$, where $a_e, b_e \geq 0$ are edge-specific constants.

To study the effect of slowdown information on self-interested behavior, this paper models the above routing problem as a heterogeneous non-atomic congestion game. The slowdown-sensitivities of the users in commodity c are modeled by a monotone, nondecreasing function $\beta^c : [0, r^c] \rightarrow [0, 1]$, where each user $x \in [0, r^c]$ has a slowdown sensitivity $\beta^{c,x} \in [0, 1]$. Here, if user x has $\beta^{c,x} = 0$, this indicates that the user is purely delay-averse and is unresponsive to slowdowns. On the other hand, if user x has $\beta^{c,x} = 1$, this indicates that the user cares nothing for *actual* delay, and selects routes solely on the basis of how over-congested the route is compared to nominal. The analysis in this paper assumes that each sensitivity distribution function β^c is unknown to the system operator, and we write \mathcal{B} to denote the set of all feasible sensitivity distribution functions

The system operator provides all users with information regarding the slowdown experienced on the various paths; this is modeled by a number $\gamma \in [0, 1]$. When $\gamma = 1$, users are provided complete and true information regarding each road's slowdown; when $\gamma = 0$, users are provided no information regarding slowdowns. Altogether, given a flow f , the subjective cost that user $x \in [0, r^c]$ experiences for using path $p \in \mathcal{P}^c$ is of the form

$$\begin{aligned} J^{c,x}(p; f) &= \sum_{e \in p} [(1 - \gamma\beta^{c,x}) \ell_e(f_e) + \gamma\beta^{c,x} (\ell_e(f_e) - \ell_e(0))] \\ &= \sum_{e \in p} [\ell_e(f_e) - \gamma\beta^{c,x} \ell_e(0)]. \end{aligned} \quad (3)$$

Thus, each user $x \in [0, r^c]$ can be viewed as interpreting the cost of a road as the difference between its latency and a moderately-scaled version of its free-flow delay. We assume that each user selects the lowest-cost path from the available

source-destination paths. We call a flow f a *Nash flow* if all users are individually using minimum-cost paths given the choices of other users. That is, for all commodities $c \in \mathcal{C}$, every user $x \in [0, r^c]$ using path p in f experiences a cost satisfying

$$J^{c,x}(p; f) = \min_{\tilde{p} \in \mathcal{P}^c} J^{c,x}(\tilde{p}; f). \quad (4)$$

The above game is an *exact potential game* with a convex potential function (see Lemma 2), which implies that its set of Nash flows is nonempty and convex.

2.2 Avoiding Perverse Signaling

As an initial step towards characterizing a system operator's optimal signaling policy, this paper compares the equilibrium total latency resulting from slowdown signaling with the un-influenced equilibrium total latency. One simple measure which captures the possible *harm* of slowdown signaling is the *perverse index* as introduced in [9]:

$$\text{PI}(\mathcal{G}, \gamma) := \sup_{G \in \mathcal{G}} \sup_{\beta \in \mathcal{B}} \frac{\mathcal{L}^{\text{nf}}(G, \beta, \gamma)}{\mathcal{L}^{\text{nf}}(G, \beta, 0)}. \quad (5)$$

Here, if a class of routing problems \mathcal{G} and signal parameter γ have a large perversity index, this indicates that there exist routing problems in \mathcal{G} for which it would be better for the system planner to avoid signaling altogether. If $\text{PI}(\mathcal{G}, \gamma) > 1$ for some γ , we say that γ is a *perverse* signaling policy. If $\text{PI}(\mathcal{G}, \gamma) = 1$, we say that γ is *non-perverse*.

3 Related Work

Since the seminal work on price of anarchy for nonatomic routing games [22], these games have provided fertile ground to investigate a wide range of questions. A common theme in these involves augmenting users' costs in some way to investigate the effect of some external influence or internal bias. For example, the altruism models of [5, 10] assume that each member $x \in [0, 1]$ of the population has an "altruism" parameter $\alpha^x \in [0, 1]$ which enters their edge cost function in the following way:

$$J_e^x(f_e) = \ell_e(f_e) + \alpha^x f_e \ell'_e(f_e). \quad (6)$$

Here, $\alpha^x = 1$ represents a fully altruistic user whose singular goal is to reduce aggregate congestion, and $\alpha^x = 0$ represents a fully selfish user whose singular goal is to minimize personal travel time. This model of altruism has been studied extensively, and much is known about the effects of this type of altruistic bias— notably, in many situations, increased levels of altruism lead to reductions in equilibrium delays. Many other similar cost function biases have been studied which take a similar form of $\ell_e(f_e) + b_e(f_e)$ for some specified $b_e(f_e)$, including pessimism [19], risk-aversion [17], and various generalizations [15, 20].

The above work is concerned with *characterizing* the effect of user preferences on equilibrium behavior. In parallel, many researchers have investigated methods of *influencing* user preferences to improve equilibrium behavior. These works on influence often involve similar modifications of user costs. For example, pricing [14] is a common means of influencing user choices, and a popular pricing model allows the system planner to assign pricing functions to each edge of $\tau_e(f_e)$, and assumes (similar to the altruism model above) that users have heterogeneous price-sensitivities s^x :

$$J_e^x(f_e) = \ell_e(f_e) + s^x \tau_e(f_e). \quad (7)$$

This is clearly tightly connected with the altruism model, particularly if the $\{\tau_e\}$ are chosen to be marginal-cost prices of the form $\tau_e^{\text{mc}}(f_e) = f_e \ell'_e(f_e)$. This has led to various synergies between the study of altruism and marginal-cost pricing [5], as results in one area neatly imply conclusions in the other.

4 Our Contributions on Slowdown Signaling

4.1 Providing Slowdown Information Can Harm Congestion on Some Networks

Our first result shows that no universal signaling policy exists: networks exist for which all nonzero slowdown signals can increase aggregate congestion costs. However, some classes of networks are immune to these pathologies; in particular, slowdown signaling has the potential to improve congestion on all parallel networks. In the following, for simplicity of exposition, each edge's latency function is assumed to be a polynomial of the form $\ell_e(f_e) = a_e(f_e)^d + b_e$, for edge-specific nonnegative coefficients a_e and b_e .

Theorem 1. *Let \mathcal{G} be the class of all routing problems. For all $\gamma \in (0, 1]$, routing problems exist for which γ is a perverse signal:*

$$\text{PI}(\mathcal{G}, \gamma) > 1. \quad (8)$$

However, let \mathcal{G}_p^d be the class of all parallel routing problems with latency functions of the form $\ell_e(f_e) = a_e(f_e)^d + b_e$. Then non-perverse signaling is possible. In particular,

$$\gamma \in [0, d/(d+1)] \quad \text{if and only if} \quad \text{PI}(\mathcal{G}_p^d, \gamma) = 1. \quad (9)$$

Note that since $\mathcal{G}_p^d \subset \mathcal{G}$, the results in (8) and (9) indicate that there exist networks for which all signaling policies are perverse, but that these pathological networks are *never* parallel networks. Thus, a system planner who knows they are working with parallel networks can safely employ any signal satisfying $\gamma \in [0, d/(d+1)]$ without fear of causing harm relative to the uninfluenced equilibria.

The proof of this theorem relies on Lemma 1, which explicitly relates the equilibrium flows under slowdown signaling to a similar formulation which is reminiscent of user costs experienced under marginal-cost tolls. This allows us to leverage existing results on marginal-cost tolls to obtain the proof of the theorem.

Lemma 1. *Let $\beta \in \mathcal{B}$, $G \in \mathcal{G}^d$, $\gamma \in [0, 1]$. For each $x \in [0, 1]$, let*

$$\alpha^x := \frac{\gamma\beta^x}{d(1 - \gamma\beta^x)}. \tag{10}$$

Then the set of Nash flows associated with the game (G, β, γ) is equal to the set of Nash flows for a game with the same network and a user-specific edge cost function of

$$J_e^x(f_e) = (1 + d\alpha^x)a_e(f_e)^d + b_e. \tag{11}$$

The proof of Lemma 1 is provided in the Appendix. The cost functions (11) in Lemma 1 are well-studied in the literature on heterogeneous nonatomic routing games. These cost functions are identical to those induced by marginal-cost tolls when users are heterogeneous in price sensitivity; they also appear in the α -altruism model of [5,10]. Thus, results from those two streams of work may be leveraged to draw conclusions about the effects of slowdown signaling.

Proof of Theorem 1 Due to Lemma 1, the proof of Theorem 1 follows in a straightforward manner from Theorem (2) of [9] and Theorem 7.1 of [10]. \square

4.2 Optimal Signaling Need Not Be Truthful

Our next result considers the goal of computing the optimal signaling policy which minimizes worst-case equilibrium traffic congestion. Here, we consider a situation in which the planner possesses additional information and in which case aggressive signaling mechanisms may be warranted. One example of this is reminiscent of the setting studied in [8]: the planner knows *a priori* that the traffic rate on the network is high enough that all edges would be used in an un-influenced Nash flow. Here, suppose that the planner knows that for every user $x \in [0, 1]$, the slowdown-sensitivity satisfies $\beta^x \in [\beta_L, \beta_U]$, where $\beta_L > 0$ and $\beta_U < 1$. In this setting, for linear-affine-cost parallel networks, the planner’s optimal signal may actually *over-state* the network’s true slowdowns:

Theorem 2. *Let $\bar{\mathcal{G}}_p^1$ denote the class of linear-latency parallel networks in which every edge has positive traffic in an un-influenced Nash flow. For every $G \in \bar{\mathcal{G}}_p^1$, it holds that*

$$\gamma^* := \frac{1}{\beta_L + \beta_U} = \arg \min_{\gamma \geq 0} \max_{\beta} \mathcal{L}^{\text{nf}}(G, \beta, \gamma). \tag{12}$$

Under the influence of this signal, the worst-case equilibrium total latency satisfies

$$\max_{G \in \bar{\mathcal{G}}_p^1} \max_{\beta} \frac{\mathcal{L}^{\text{nf}}(G, \beta, \gamma^*)}{\mathcal{L}^*(G)} = \frac{4}{3} \left(1 - \frac{\beta_L/\beta_U}{(1 + \beta_L/\beta_U)^2} \right). \tag{13}$$

The proof of Theorem 2 is provided in the Appendix.

4.3 Broader Implications for Heterogeneous Congestion Games

An interesting byproduct of the analysis required for this work is that we have discovered a weighted potential game formulation which applies to a large class of *heterogeneous* nonatomic congestion games; to the best of our knowledge, this formulation is novel. This formulation opens the door to computing Nash flows efficiently for a large class of games by performing gradient descent on the potential function of an associated exact potential game. To understand the new formulation, first consider the following lemma which demonstrates that the slowdown-sensitivity games in this paper are exact potential games:

Lemma 2. *For any $\beta \in \mathcal{B}$, $G \in \mathcal{G}$, and $\gamma \geq 0$, the game (G, β, γ) specified by cost functions (3) is an exact potential game with a convex potential function. Thus:*

1. (G, β, γ) has at least one Nash flow,
2. the set of Nash flows is convex, and
3. for each user $x \in [0, 1]$, given two Nash flows f and f' (in which user x chooses paths p and p' , respectively), it holds that $J^x(p; f) = J^x(p'; f')$.

The proof of Lemma 2 is provided in the Appendix. We can now state the main result, that many heterogeneous nonatomic congestion games with a particular form of polynomial cost function are weighted potential games and that their Nash flows exactly coincide with those of our slowdown-sensitivity games.

Theorem 3. *In some routing problem $G \in \mathcal{G}^d$, let every user $x \in [0, 1]$ have altruism parameter $\alpha^x \in [0, 1]$ and experience a cost on edge e of $J_e^x(f) = (1 + d\alpha^x)a_e(f_e)^d + b_e$. Then routing problem G coupled with the user population described by α^x is a weighted potential game whose set of Nash flows is equal to that of slowdown-sensitivity game $(G, \beta, 1)$, where for each $x \in [0, 1]$ it holds that*

$$\beta^x = \frac{\alpha^x d}{\alpha^x d + 1}. \tag{14}$$

The proof of Theorem 3 is provided in the Appendix.

5 Conclusion

This paper explores the connections between a new class of slowdown-sensitive congestion games and established results for heterogeneous nonatomic routing games under the influence of altruism and marginal-cost pricing. In particular, we demonstrate a tight connection between the two, and illustrate how to connect results from one to the other. Finally, we exploit this connection to demonstrate a novel result for certain heterogeneous nonatomic congestion games, namely that they are weighted potential games and thus their equilibria may be computed efficiently.

Appendix: Proofs

Proof of Lemma 1. By (10), emulated cost functions (11) can be written

$$J_e^x(f_e) = \frac{a_e(f_e)^d}{1 - \gamma\beta^x} + b_e. \tag{15}$$

Since users’ ordinal preferences (and thus Nash flows) are invariant to multiplication by user-specific constants, the cost functions in (15) are equivalent to ones given by

$$\tilde{J}_e^x(f_e) = a_e(f_e)^d + (1 - \gamma\beta^x)b_e. \tag{16}$$

Evidently, the cost functions (16) are equal to those of the nominal game (G, β, γ) (see (3)). Since, for each user x , these cost functions encode the same ordinal preferences as those given by (11), the two sets of cost functions induce identical sets of Nash flows. \square

Proof of Theorem 2. The optimal signal factor (12) can be deduced from the results in [8, Theorem 1] regarding the optimal marginal-cost-toll scale factor in an identical setting. Our Lemma 1 provides that once a value of γ is fixed, each slowdown-sensitivity β^x distribution can be used to compute a taxation-sensitivity α^x distribution which induces an identical set of Nash flows. In [8, Theorem 1], it is shown that worst-case congestion is minimized for scaled marginal-cost taxes when the taxation-sensitivity distribution satisfies

$$\max_x \alpha^x = \frac{1}{\min_x \alpha^x}. \tag{17}$$

Thus, due to the equivalence between slowdown-sensitivity and taxation sensitivity and applying the transformation (10) from Lemma 1, the optimal signal factor γ^* satisfies

$$\frac{\gamma^* \beta_L}{1 - \gamma^* \beta_L} = \frac{\gamma^* \beta_U}{1 - \gamma^* \beta_U}, \tag{18}$$

which implies (12). The upper bound (13) follows immediately from [6, Lemma 3.1]. \square

Proof of Lemma 2. Note that the individual slowdown sensitive cost function (3) is a sum of two terms: $\ell_e(f_e)$ and $\gamma\beta^x\ell_e(0)$. A game with costs defined by only the first term is simply a homogeneous nonatomic routing game, which is well-known to be an exact potential game with a convex potential function [22,23]. A game with costs defined by only the second term is a trivial game in which each user’s cost function depends only on her own action; any such game is also known to be an exact potential game with a linear potential function [23]. In [23] it is shown that a game whose cost functions are the sum of exact potential games’ cost functions is itself an exact potential game whose potential function is the sum of the potential functions of its component games. Therefore, (G, β, γ) is an exact potential game with a convex potential function—and this convexity implies points (1), (2), and (3). \square

Proof of Theorem 3. The proof relies on a serial application of Lemmas 1 and 2. In particular, Lemma 1 provides a bijection between the heterogeneous altruistic congestion games assumed by Theorem 3 and the slowdown-sensitivity games considered in this paper. Equation (14) is simply the inverse mapping of (10); thus, the cost functions of the two games express identical ordinal preferences and are related by a set of user-specific constant multipliers. Thus, the games are (for the purposes of equilibrium computation) equivalent: equilibria computed for one are automatically equilibria for the other. Accordingly, in light of Lemma 2, one may use potential game equilibrium-finding techniques (e.g., gradient descent) to compute the equilibria of the slowdown-sensitivity game and thereby find the equilibria of the altruistic congestion game. \square

References

1. Acemoglu, D., Makhdoumi, A., Malekian, A., Ozdaglar, A.: Informational Braess' paradox: the effect of information on traffic congestion. *Oper. Res.* **66**(4), 893–917 (2018)
2. Beckmann, M.J., McGuire, C.B., Winsten, C.B.: *Studies in the Economics of Transportation* (1955). <https://doi.org/10.1057/jors.1980.83>
3. Biyik, E., Lazar, D., Pedarsani, R., Sadigh, D.: Incentivizing efficient equilibria in traffic networks with mixed autonomy. *IEEE Trans. Control Netw. Syst.* **8**(4), 1717–1729 (2021). <https://doi.org/10.1109/TCNS.2021.3084045>, <https://ieeexplore.ieee.org/document/9442301/>
4. Bonifaci, V., Salek, M., Schäfer, G.: Efficiency of restricted tolls in non-atomic network routing games. In: Persiano, G. (ed.) *SAGT 2011*. LNCS, vol. 6982, pp. 302–313. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-24829-0_27
5. Brown, P.N.: When altruism is worse than anarchy in nonatomic congestion games. In: *2021 American Control Conference (ACC)*, pp. 4503–4508. IEEE (2021). <https://doi.org/10.23919/ACC50511.2021.9483340>, <https://arxiv.org/abs/2007.05591>, <https://ieeexplore.ieee.org/document/9483340/>
6. Brown, P.N., Marden, J.R.: A study on price-discrimination for robust social coordination. In: *American Control Conference*, Boston, MA, USA, pp. 1699–1704 (2016)
7. Brown, P.N., Marden, J.R.: Studies on robust social influence mechanisms: incentives for efficient network routing in uncertain settings. *IEEE Control Syst. Mag.* **37**(1), 98–115 (2017). <https://doi.org/10.1109/MCS.2016.2621482>
8. Brown, P.N., Marden, J.R.: The robustness of marginal-cost taxes in affine congestion games. *Trans. Autom. Control* **62**(8), 3999–4004 (2017)
9. Brown, P.N., Marden, J.R.: Can taxes improve congestion on all networks? *IEEE Trans. Control Netw. Syst.* **7**(4), 1643–1653 (2020). <https://doi.org/10.1109/TCNS.2020.2992679>
10. Chen, P.a., De Keijzer, B., Kempe, D., Shaefer, G.: Altruism and its impact on the price of anarchy. *ACM Trans. Econ. Comput.* **2**(4), 1–45 (2014)
11. Cheng, Y., Langbort, C.: A model of informational nudging in transportation networks. In: *2016 IEEE 55th Conference on Decision and Control, CDC 2016*, pp. 7598–7604 (2016). <https://doi.org/10.1109/CDC.2016.7799443>

12. Cole, R., Dodis, Y., Roughgarden, T.: Pricing network edges for heterogeneous selfish users. In: Proceedings of the 35th ACM Symposium on Theory of Computing, New York, pp. 521–530 (2003). <https://doi.org/10.1145/780615.780618>, <https://portal.acm.org/citation.cfm?doid=780542.780618>
13. Colini-Baldeschi, R., Klimm, M., Scarsini, M.: Demand-Independent Tolls, pp. 1–18 (2017). <https://arxiv.org/abs/1708.02737>
14. Ferguson, B.L., Brown, P.N., Marden, J.R.: The effectiveness of subsidies and tolls in congestion games. *IEEE Trans. Autom. Control* **67**(6), 2729–2742 (2022). <https://doi.org/10.1109/TAC.2021.3088412>, <https://arxiv.org/abs/2102.09655>, <https://ieeexplore.ieee.org/document/9451652/>
15. Kleer, P., Schäfer, G.: Path deviations outperform approximate stability in heterogeneous congestion games. In: International Symposium on Algorithmic Game Theory, pp. 212–224 (2017). <https://arxiv.org/abs/1707.01278>
16. Li, R., Brown, P.N., Horowitz, R.: Employing altruistic vehicles at on-ramps to improve the social traffic conditions. In: 2021 American Control Conference (ACC), pp. 4547–4552. IEEE (2021). <https://doi.org/10.23919/ACC50511.2021.9482993>, <https://ieeexplore.ieee.org/document/9482993/>
17. Lianas, T., Nikolova, E., Stier-Moses, N.E.: Asymptotically tight bounds for inefficiency in risk-averse selfish routing. In: IJCAI International Joint Conference on Artificial Intelligence, vol. 2016, pp. 338–344 (2016)
18. Massicot, O., Langbort, C.: Public signals and persuasion for road network congestion games under vagaries. *IFAC-PapersOnLine* **51**(34), 124–130 (2019)
19. Meir, R., Parkes, D.C.: Playing the wrong game: smoothness bounds for congestion games with behavioral biases. *Perform. Eval. Rev.* **43**(3), 67–70 (2015)
20. Meir, R., Parkes, D.C.: Playing the wrong game: bounding negative externalities in diverse populations of agents. In: 17th International Conference on Autonomous Agents and MultiAgent Systems, pp. 86–94 (2018). <https://arxiv.org/pdf/1411.1751.pdf>
21. Pigou, A.C.: *The Economics of Welfare*. Macmillan, New York (1920)
22. Roughgarden, T.: *Selfish Routing and the Price of Anarchy*. MIT Press, Cambridge (2005)
23. Sandholm, W.H.: *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge (2009)
24. Thaler, R.H., Sunstein, C.R.: *Nudge* (2008)
25. Tversky, A., Kahneman, D.: *Judgment under Uncertainty: Heuristics and Biases*. Science (1974)



Budgeted Adversarial Network Resource Utilization Games

Yi Zhang and Sanjiv Kapoor^(✉)

Illinois Institute of Technology, Chicago, IL 60616, USA
yzhan257@hawk.iit.edu, kapoor@iit.edu

Abstract. This paper studies budgeted adversarial resource utilization game, where one of the player's (designer) strategy is the utilization of resources while the other player's (adversary) role is to police the resources for misuse. In this context, we consider routing games where a designer plans routes on a computer network and the adversary intercepts the routes on the network. Another example is in determining adversarial strategies to block access to travel or resources that may be considered to pose a risk to society, e.g. during a pandemic where the population (designer) goal may not be coincide with the societal goal of minimizing accessing a banned resource. We model this as a zero-sum game with constraints on the adversary or designer budgets. While zero-sum games can be solved using linear programs, we illustrate faster combinatorial methods to solve the problem. We first consider the resource access problem game on a bipartite graph where both the designer and the adversary have independent budget constraints and distinct costs and show a fast algorithm to determine a Nash equilibrium. We also consider the situation where the designer would strategize on paths in a general graph. In this application of determining network paths, where the adversary would attack edges in order to block the paths, we also discuss the case of multiple designers and, in particular show faster algorithms when there are 2 designers. These results utilize properties of minimum cuts in 2-commodity flow routing.

Keywords: Security · Network · Game theory · Nash equilibrium

1 Introduction

In this paper we consider a budgeted adversarial resource utilization game, where a designer plans the utilization of resources when facing an adversary. The adversary is a user who attempts to prevent the designer from using resources. The adversary and/or the designer have budgets. Typical examples would include public health restrictions on movements during pandemics (where analysis of behavior and policies using game theory has been investigated [4, 11] or ensuring network security. While no moral subjective is assigned to the players, a rational approach of a traveller would be to adopt a route which has the least stringent

governmental restrictions when planning airline trips across boundaries, while inadvertently carrying the virus. The governing public health policy takes on the role of an adversary with tests and mandates within the travel network, placing restrictions on edges to prevent virus transmission through human or even goods transportation.

Security issues in resource allocations games have been considered extensively. Applications in communication and transportation networks have been studied via game theoretic analysis in [3, 12, 15, 16]. An example problem is securing the connectivity in communication systems on a graph where a set of nodes is to be connected using a spanning tree. The resources are edges and a subset of edges provides utility to the designer. The adversary may attack one or more edges, each edge having a cost of attack. The adversary has a budget which cannot be exceeded.

This paper considers network security games in the larger context of resource utilization adversarial games. Given a set of resource elements, a collection of resource subsets defines designer strategies. Each resource element has a cost of attack and the adversary has a budget. This is a zero-sum game and it is required to determine a Nash equilibrium in this model with strategies over a continuous range. Pure Nash equilibrium with integral feasible strategies, where attacker pays for entire cost of edges if present in strategy set, need not always exist. Mixed equilibrium over the integral strategy sets exists but additional budget constraints impact the optimality. Budget utilization leading to fractional solutions provides for challenging problems, while mixed solutions over finite integral strategies is not considered and left for further work.

In this paper consider an important application of this framework, a network routing game where a designer plans routes (paths) on a network and the adversary intercepts the routes on the network. The designer has to route paths between a source-destination pair and can choose multiple paths to satisfy her unit requirement. The fractional solution can also be interpreted as the probability of the designer choosing the particular path. The adversary chooses edges to attack simultaneously. Each attack on an edge incurs a cost. The attack on edges is fractional and may contain multiple edges as long as the budget allows. As indicated above, the role of budget is particularly interesting and there exists an example where paying for the entire edge does not result in a Nash equilibrium. The role of the designer and adversary may be reversed as in the case of drug-trafficking [19].

Computing Nash equilibrium in general games is a hard problem, even for mixed strategies. The problem is PPAD-hard [5, 14] and polynomial solutions are not known. An important sub-class is the class of zero-sum games. Zero-sum games are easier to compute using linear programming formulation and require time polynomial in the size of strategy space. However, in most cases the strategy space is exponential in the description of the problem, as in the case of the routing or connectivity games. To obtain polynomial solutions, one approach is to utilize separation oracles and the ellipsoid method [6] for solving linear programs, which unfortunately has substantial complexity, albeit polynomial. Furthermore, additional budget constraints on the strategies adds additional

complication and may require enumeration of the feasible sets that form the support set. An approach to tackle the exponential space is to employ duality theory to identify problem characteristics.

In this paper we consider budgeted versions of zero-sum games and describe polynomial method to solve these games. While the complexity of zero-sum games with succinct representation has been considered, e.g. for Colonel Blotto games [1] and the ellipsoid method is used, the budgeted case is not addressed in that research. Costs on elements have been addressed in the case of matroidal zero-sum games where the defender or designer chooses a basis B of the matroid and the adversary incurs a cost of attack on elements of the subsets [2]. That research considers a utility function that incorporates the costs without any explicit budget constraint. Additionally, research [8] has also considered the problem of designing a spanning tree and attacks on the tree using a single edge attack. There is a cost of attack on edges which, again, is incorporated into the objective and the paper describes how to identify a critical set of edges over which a mixed equilibrium can be determined. Addition of budgets has direct impact on the feasible polyhedron describing the strategy space and thus complicates the problem. Related security games have been studied by [20], where the adversary only picks one target to attack. Additionally, research in [17] [18] extend the model to incorporate multiple attack targets and to target network security; however targets or elements are not distinguished by differing costs or weights. Arguably, adversaries typically have budgets within which they attempt to inflict as much damage on resources, say a communication network, and our current approach models real life scenarios effectively. Budgeted attacks have also been considered in [7] where constrained optimization methods are used to determine the expected benefits of the adversary and defender. However, no specific form of the underlying problem are used and the algorithms are not guaranteed to be polynomial.

The key challenge in the analysis of designer-adversary strategies is the exponential number of strategies, even when the game can be described succinctly. As we illustrate, it is possible to solve classes of these problems using combinatorial interpretations based on primal-dual formulations. The optimization programs for the designer and attacker are constrained and interdependent. The key technique we employ is to combine both the defender's and adversary's primal and dual constraints in a complementary fashion to formulate a composite linear program and its dual. We then apply a primal-dual algorithm that is shown to be efficient. We first illustrate this for the case of resource access bipartite networks with independent designer and adversary costs and budgets. Additionally for our general network application problem, the routing problem, we utilize primal-dual programs to characterize the Nash equilibrium solutions. The solutions based on these characterizations can be obtained by simple combinatorial max-flow algorithms that also identify the minimum cut in a graph. In particular we also consider the case when there are multiple designers and provide combinatorial solutions for the case of 2-designers. This is a particularly interesting case as 2-commodity network flows have interesting cut theorems and combinatorial algorithms exist for this problem [9, 10].

2 Preliminaries and Results

2.1 Problem Definitions and Notations

We first define the games that we will consider in this paper.

Bipartite Routing Adversarial Game: In the *Bipartite Routing Adversarial Game* we consider a bipartite network $N(E, d, c)$. $E = \{e_1, e_2, \dots, e_m\}$ is a bipartition of m edges, where each edge e_i is associated with a positive design cost $d[i]$ and attack cost $c[i]$. $d, c \in \mathbb{R}_+^m$ are the cost vectors. We denote the designer's and the adversary's budget by B_D and B_A , respectively.

The flow design is represented by a vector $f \in [0, 1]^m$ with the requirement $\sum_{i=1}^m f[i] = 1$, where $f[i]$ is the flow amount on edge e_i . For each edge e_i with flow amount $f[i]$, the designer's cost is $d[i]f[i]$. We say a design f is within budget if $\sum_{i=1}^m d[i]f[i] \leq B_D$.

An attack is a vector $X \in [0, 1]^m$ indicating the attack on edges, where $X[i]$ denotes the attack, or attack level, on edge e_i , the fraction of flow captured on e_i . We say an attack X is within budget if $\sum_{i=1}^m c[i]X[i] \leq B_A$.

The adversary's benefit is the sum of flow she captures over the bipartition and defined as $\sum_{j=1}^m f[j] \cdot X[j]$. This is a zero-sum game where the adversary maximizes her benefit by choosing an attack strategy and the designer minimizes it by choosing a flow design.

Network Routing Adversarial Game: In the *Network Routing Adversarial Game*, we consider a network $N(G, s, t, c)$ where $G = (V, E)$ is an undirected graph, V is a set of vertices with $|V| = n$ and $E = \{e_1, e_2, \dots, e_m\}$ is a set of m edges, where each edge e_i is associated with a positive adversary cost $c[i]$. The adversary has a budget $B > 0$. The source destination pair is $(s, t) \in V$. The flow designer is required to determine paths to route a normalized unit of traffic/flow over $\mathcal{P} = \{p_1, p_2, \dots, p_k\}$, the collection of all k possible $s - t$ paths. In the case of r -commodity network, each designer i has her source and sink pair s_i, t_i and her set of path \mathcal{P}_i , where $|\mathcal{P}_i| = k_i$. The overall path set is denoted by \mathcal{P} , with $|\mathcal{P}| = k$ where $k = \sum_{i=1}^r k_i$.

The flow design is represented by a vector $f \in [0, 1]^k$ with the requirement $\sum_{j=1}^k f[j] = 1$, where $f[j]$ is the flow amount on path p_j . We denote by $f_e[i] = \sum_{j|p_j \ni e_i} f[j]$ the flow on edge e_i , which is the summation of flow on every path containing e_i . In the case of multiple designers, $f_e[i]$ includes the flow of all designers.

An attack is a vector $X \in [0, 1]^m$ indicating the attack on edges, where $X[i]$ denotes the attack, or attack level, on edge e_i . We say an attack X is within budget if $\sum_{i=1}^m c[i]X[i] \leq B$.

Given an attack $X \in [0, 1]^m$, the impact vector of the attack over all paths is denoted by $X_{\mathcal{P}}$ where $X_{\mathcal{P}}[j] = \min\{\sum_{i|e_i \in p_j} X[i], 1\}$. $X_{\mathcal{P}}[j]$ is the attack on a path p_j , the fraction of flow on that path captured by the adversary, which is the summation of attacks on edges on that path, up to 1.

The adversary's benefit is the sum of fractions of flow on the paths she attacks: $\sum_{j=1}^k f[j] \cdot X_{\mathcal{P}}[j]$. This is a zero-sum game where the adversary maximizes her benefit by choosing an attack strategy and the designer minimizes it by choosing a flow design.

An $s - t$ cut C is a set of edges whose removal disconnects all paths from s to t and partitions V into 2 component S, T , where $S \cap T = \emptyset$, $s \in S$, $t \in T$. We define the cost of C to be $|C| = \sum_{e_i \in C} c[i]$. Given an input graph, let C_{min} be the min-cost $s - t$ cut.

2.2 Results

In this paper we obtain the following results:

1. Our first result is a combinatorial algorithm to solve the Bipartite Routing Adversarial Game with both designer and adversary budget, modeled via a bipartite graph, in $O(m \log m)$ steps. Note that the costs of edges incurred by the designer and the adversary (attacker) are different.
2. We design a combinatorial algorithm to solve the Network Routing Adversarial game with adversary budget constraint in $O(nm)$ steps.
3. We illustrate combinatorial methods for the 2-commodity game in the context of r -commodity Network Routing Adversary game with budgets. The time complexity is $O(n^3)$.

These methods have complexities that improve the standard linear programming approach. Some of the proofs are provided in the appendix.

3 Bipartite Routing Adversarial Game

In this section we look at the **Bipartite Routing Adversarial Game**. Given a bipartite network $N(E, d, c)$ where each edge $e_i \in E, \forall i = 1 \dots m$ has a cost for designer $d[i]$ and cost for adversary $c[i]$. The designer and adversary has budget B_D, B_A respectively. We determine the designer’s strategy f and the adversary’s strategy X . We assume each adversary cost $c[i]$ and each designer cost $d[i]$ are unique and $d[i] \geq 1, \forall i$. We sort E in increasing order of $d[i]$ in $O(m \log m)$ steps if they are not sorted. In our method the designer will consider this sorted order of edges whenever her strategy requires it to add additional edges. We assume $B_D \geq d[1]$, the smallest designer cost, to ensure feasible flow design. We also assume $B_A \leq \sum_{i=1}^m c[i]$ for otherwise the adversary always captures the entire flow by fully attacking every edge.

Given the adversary’s strategy X , the designer solves the following linear program with its dual:

$$\begin{aligned}
 \text{LP1 : } \min \quad & \sum_{i=1}^m X[i]f[i] \\
 \text{s.t.} \quad & \sum_{i=1}^m f[i] \geq 1 \\
 & \sum_{i=1}^m d[i]f[i] \leq B_D
 \end{aligned} \tag{1}$$

$$\begin{aligned}
 \text{LP2: } \max \quad & X_D - B_D \Omega \\
 \text{s.t.} \quad & X_D - X[i] \leq d[i]\Omega, \forall i
 \end{aligned} \tag{2}$$

Given the designer’s strategy f , the adversary solves the following linear program with its dual:

$$\begin{aligned}
 \mathbf{LP3:} \max \sum_{i=1}^m f[i]X[i] & & \mathbf{LP4:} \min B_A\gamma + \sum_{i=1}^m \delta[i] & (4) \\
 \text{s.t.} \sum_{i=1}^m c[i]X[i] \leq B_A & (3) & \text{s.t.} c[i]\gamma + \delta[i] \geq f[i], \forall i & \\
 X[i] \leq 1, \forall i & & &
 \end{aligned}$$

Solving **LP1** and **LP3** jointly requires non-linear optimization. We avoid this by presenting an equivalent primal linear program **LP5** and its dual **LP6** to compute the Nash equilibrium with **LP5** optimizing the adversary’s strategy and **LP6** computing the optimal designer’s strategy:

$$\begin{aligned}
 \mathbf{LP5:} \min B_A\gamma + \sum_{i=1}^m \delta[i] & & \mathbf{LP6:} \max X_D - B_D\Omega & \\
 \text{s.t.} c[i]\gamma + \delta[i] \geq f[i], \forall i & & \text{s.t.} X_D - X[i] \leq d[i]\Omega, \forall i & \\
 \sum_{i=1}^m f[i] \geq 1 & (5) & \sum_{i=1}^m c[i]X[i] \leq B_A & (6) \\
 \sum_{i=1}^m d[i]f[i] \leq B_D & & X[i] \leq 1, \forall i &
 \end{aligned}$$

The optimum solution of **LP5**, **LP6** must satisfy the following complementary slackness conditions:

$$\left\{ \begin{array}{ll}
 (c[i]\gamma + \delta[i] - f[i])X[i] = 0, \forall i & \dots(i) \\
 (\sum_{i=1}^m f[i] - 1)X_D = 0 & \dots(ii) \\
 (\sum_{i=1}^m d[i]f[i] - B_D)\Omega = 0 & \dots(iii) \\
 (X_D - X[i] - d[i]\Omega)f[i] = 0, \forall i & \dots(iv) \\
 (\sum_{i=1}^m c[i]X[i] - B_A)\gamma = 0 & \dots(v) \\
 (X[i] - 1)\delta[i] = 0, \forall i & \dots(vi)
 \end{array} \right.$$

We present an algorithm to calculate a feasible solution that satisfies all the complementary slackness conditions. We observe that the structure of the Nash equilibrium evolves step by step as the designer budget B_D and opponents budget B_A increases. First assume that the attack budget B_A is very small in

which case the attack on no edge will reach 1. Denote $C(k) = \sum_{i=1}^k c[i]$ and $CD(k) = \sum_{i=1}^k c[i]d[i], \forall k = 1 \dots m$ for convenience. They can be computed in $O(m)$ steps. While the designer's goal is to even out the bang-per-buck of every edge, note that with the limited B_D she may not be able to do so on the entire cut. If the designer limits attention to only the k cheapest edges, we can calculate the flow on those edges with the same bang-per-buck γ along with the corresponding designer budget $B_D[k]$ by solving the following system of equations.

$$\left\{ \begin{array}{l} \sum_{i=1}^k f[i] = 1 \\ \sum_{i=1}^k d[i]f[i] = B_D[k] \\ f[i] = c[i]\gamma, \forall i = 1 \dots k \end{array} \right. \implies \left\{ \begin{array}{l} \gamma = \frac{1}{C(k)} \\ f[i] = \frac{c[i]}{C(k)}, \forall i = 1 \dots k \\ B_D[k] = \frac{CD(k)}{C(k)} \end{array} \right. \quad (7)$$

One can verify that $0 \leq f[i] \leq 1, \forall i = 1 \dots k$. Each $B_D[k]$ can be calculated in $O(1)$ steps. If $B_D[k] < B_D < B_D[k + 1]$, the designer's flow pattern is a mix of flow patterns corresponding to $B_D[k]$ and $B_D[k + 1]$ respectively.

Lemma 1. $B_D[k - 1] < B_D[k]$

Since $B_D[k]$ increases monotonically w.r.t. k we get:

Corollary 1. *Let \hat{k} be the largest integer with $B_D[\hat{k} - 1] < B_D$. The index \hat{k} can be found in $O(\log m)$ steps.*

The index \hat{k} is very important in our methodology. With a small B_A , the adversary's attack will focus only on the first $\hat{k} - 1$ edges. As B_A increases, the designer will have incentive to route flow on the edge with the next cheapest design cost, and the adversary's response also includes one more edge. For $t = \hat{k} \dots m$, we define $B_A[t]$ to be the highest adversary budget where e_t is not attacked. When B_A goes above $B_A[t]$ the designer will start routing on e_{t+1} and the adversary will start attacking e_t . Calculate $B_A[t]$ by solving the following system of equations:

$$\left\{ \begin{array}{l} \sum_{i=1}^t c[i]X[i] = B_A[t] \\ X[1] = 1 \\ X[t] = 0 \\ X[i] = X_D - d[i]\Omega, \forall i = 1 \dots t \end{array} \right. \implies \left\{ \begin{array}{l} X_D = \frac{d[t]}{d[t] - d[1]} \\ \Omega = \frac{1}{d[t] - d[1]} \\ X[i] = \frac{d[t] - d[i]}{d[t] - d[1]}, \forall i = 2 \dots t \\ B_A[t] = \frac{d[t]C(t) - CD(t)}{d[t] - d[1]} \end{array} \right. \quad (8)$$

One can verify that $0 \leq X[i] \leq 1, \forall i = 1 \dots t$. Each $B_A[t]$ can be calculated in $O(1)$ steps.

Lemma 2. $B_A[t - 1] < B_A[t]$.

We now construct best responses for both parties. For convenience all the unmentioned variables $X[t + 1] \dots X[m]$ take value of 0. Depending on the value of B_A , there are three cases. In each case the design and attack follow a specific pattern. In the first case where $B_A \leq B_A[\hat{k}]$, the designer routes flow on the first \hat{k} edges, the adversary budget is enough to cover only the first $\hat{k} - 1$ edges and no edge is attacked fully. In the second case where $B_A[t] < B_A \leq B_A[t + 1]$, $\hat{k} \leq t \leq m - 1$, the design includes the first t edges, the attack includes the first $t - 1$ edges and e_1 is always attacked fully. In the third case where $B_A > B_A[m]$, the flow design and the attack both uses all m edges. The Nash equilibrium in each of the cases can be computed in $O(m)$ steps.

Case(I): $B_A \leq B_A[\hat{k}]$

In this case the design uses the first \hat{k} edges and the attack uses the first $\hat{k} - 1$ edges. The attack on e_1 has not reached 1 yet. We first calculate the designer's best response f^* along with the dual variables by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^{\hat{k}} f^*[i] = 1 \\ \sum_{i=1}^{\hat{k}} f^*[i]d[i] = B_D \\ c[i]\gamma = f^*[i], \forall i = 1 \dots \hat{k} - 1 \end{array} \right. \implies \left\{ \begin{array}{l} \gamma = \frac{d[\hat{k}] - B_D}{d[\hat{k}] \cdot C(\hat{k} - 1) - CD(\hat{k} - 1)} \\ f^*[\hat{k}] = \frac{B_D \cdot C(\hat{k} - 1) - CD(\hat{k} - 1)}{d[\hat{k}] \cdot C(\hat{k} - 1) - CD(\hat{k} - 1)} \\ f^*[i] = \frac{c[i](d[\hat{k}] - B_D)}{d[\hat{k}] \cdot C(\hat{k} - 1) - CD(\hat{k} - 1)}, \\ \forall i = 1 \dots \hat{k} - 1 \end{array} \right.$$

Lemma 3. f^*, γ are feasible solutions to the linear program.

Now we calculate the adversary's best response X^* with dual variables X_D, Ω by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^{\hat{k}} c[i]X^*[i] = B_A \\ X^*[\hat{k}] = 0 \\ X_D - X^*[i] = d[i]\Omega, \forall i = 1 \dots \hat{k} \end{array} \right. \implies \left\{ \begin{array}{l} X_D = \frac{d[\hat{k}]B_A}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})} \\ \Omega = \frac{B_A}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})} \\ X^*[i] = \frac{(d[\hat{k}] - d[i])B_A}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})}, \forall i = 1 \dots \hat{k} \end{array} \right.$$

Lemma 4. X^*, X_D, Ω are feasible solutions to the linear program.

Lemma 5. The best responses calculated above form a Nash equilibrium.

Case(II): $B_A[t] < B_A \leq B_A[t + 1]$, $\hat{k} \leq t \leq m - 1$

In this case the design uses the first t edges and the attack uses the first $t - 1$ edges. The attack on e_1 is always 1. Recall that $B_A[t]$ increases monotonically w.r.t. t . Therefore we can search for the value of t in $O(\log m)$ steps. We now

calculate the designer's best response f^* alone with the dual variables by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^t f^*[i] = 1 \\ \sum_{i=1}^t f^*[i]d[i] = B_D \\ c[1]\gamma + \delta[1] = f^*[1] \\ c[i]\gamma = f^*[i], \forall i = 2 \dots t \end{array} \right\} \implies \left\{ \begin{array}{l} \gamma = \frac{B_D - d[1]}{CD(t) - d[1]C(t)} \\ \delta[1] = \frac{CD(t) - B_D C(t)}{CD(t) - d[1]C(t)} \\ f^*[1] = \frac{c[1](B_D - d[1]) + CD(t) - B_D C(t)}{CD(t) - d[1]C(t)} \\ f^*[i] = \frac{c[i](B_D - d[1])}{CD(t) - d[1]C(t)}, \forall i = 2 \dots t \end{array} \right.$$

Lemma 6. f^*, δ, γ are feasible solutions to the linear program.

Now we calculate the adversary's best response X^* with dual variables X_D, Ω by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^t c[i]X^*[i] = B_A \\ X^*[1] = 1 \\ X_D - X^*[i] = d[i]\Omega, \forall i = 1 \dots t \end{array} \right\} \implies \left\{ \begin{array}{l} X_D = \frac{CD(t) - d[1]B_A}{CD(t) - d[1]C(t)} \\ \Omega = \frac{C(t) - B_A}{CD(t) - d[1]C(t)} \\ X^*[i] = \frac{CD(t) - d[1]B_A - d[i](C(t) - B_A)}{CD(t) - d[1]C(t)}, \\ \forall i = 1 \dots t \end{array} \right.$$

Lemma 7. X^*, X_D, Ω are feasible solutions to the linear program.

Lemma 8. The best responses calculated above form a Nash equilibrium.

Case(III): $B_A > B_A[m]$

In this case both the design and the attack uses all m edges and the attack on e_1 is 1. As B_A increases the attack become closer and closer to a full attack on every edge. We first calculate the designer's best response f^* alone with the dual variables by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^m f^*[i] = 1 \\ \sum_{i=1}^m f^*[i]d[i] = B_D \\ c[1]\gamma + \delta[1] = f^*[1] \\ c[i]\gamma = f^*[i], \forall i = 2 \dots m \end{array} \right\} \implies \left\{ \begin{array}{l} \gamma = \frac{B_D - d[1]}{CD(m) - d[1]C(m)} \\ \delta[1] = \frac{CD(m) - B_D \cdot C(m)}{CD(m) - d[1]C(m)} \\ f^*[1] = \frac{c[1](B_D - d[1]) + CD(m) - B_D \cdot C(m)}{CD(m) - d[1]C(m)} \\ f^*[i] = \frac{c[i](B_D - d[1])}{CD(m) - d[1]C(m)}, \forall i = 2 \dots m \end{array} \right.$$

Lemma 9. f^*, δ, γ are feasible solutions to the linear program.

Now we calculate the adversary's best response X^* with dual variables X_D, Ω by solving the following system of equations in $O(m)$ steps.

$$\left\{ \begin{array}{l} \sum_{i=1}^m c[i]X^*[i] = B_A \\ X^*[1] = 1 \\ X_D - X^*[i] = d[i]\Omega, \forall i = 1 \dots m \end{array} \right\} \implies \left\{ \begin{array}{l} X_D = \frac{CD(m) - d[1]B_A}{CD(m) - d[1]C(m)} \\ \Omega = \frac{C(m) - B_A}{CD(m) - d[1]C(m)} \\ X^*[i] = \frac{CD(m) - d[1]B_A - d[i](C(m) - B_A)}{CD(m) - d[1]C(m)}, \\ \forall i = 1 \dots m \end{array} \right.$$

Lemma 10. X^*, X_D, Ω are feasible solutions to the linear program.

Lemma 11. The best responses calculated above form a Nash equilibrium.

To summarize, here is the algorithm to compute the Nash equilibrium of **Bipartite Routing Adversarial Game**.

Algorithm 1. Bipartite Routing Adversarial Game

Input: $N(E, d, c), B_D, B_A$

Output: f^*, X^*

- 1: Sort d in increasing order if not already sorted
 - 2: Calculate $C(i), CD(i), \forall i = 1 \dots m$
 - 3: Calculate $B_D[i], \forall i = 1 \dots m$ using **Equation(7)**
 - 4: Determine \hat{k} using **Corollary 1**
 - 5: Calculate $B_A[i], \forall i = \hat{k} \dots m$ using **Equation(8)**
 - 6: Determine the case from one of **Case(I),(II),(III)** using B_A and $B_A[i]$ and calculate the best responses f^*, X^*
-

Theorem 1. A Nash equilibrium in a **Bipartite Routing Adversarial Game** can be computed in $O(m \log m)$ steps. With pre-sorted designer costs $d[i]$, the Nash equilibrium can be computed in $O(m)$ steps.

4 Network Routing Adversarial Game

In this section we look at the **Network Routing Adversarial Game**. Given a network $N(G, s, t, c)$, let E be the edge set and \mathcal{P} be the set of all $s - t$ paths. Each edge e_i is associated with an adversary cost $c[i]$ and the adversary has a budget B . Note that in this version of game there is no budget constraint for the designer.

In order to solve this zero-sum game we establish linear programs whose solution provides the routing and attack strategy. We note that this is a non-atomic routing game, not a mixed equilibrium game, though it may be interpreted as one.

Since the number of possible paths in the network is exponential, we would attempt to utilize combinatorial methods. We first consider the adversary’s strategy. Let the designer’s flow pattern be f ; then the adversary solves the following program with its dual:

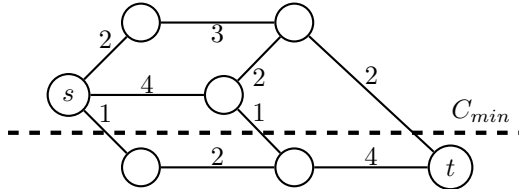
$$\begin{aligned}
 \text{LP9: } \max \quad & \sum_{j=1}^k f[j] \cdot X_{\mathcal{P}}[j] \\
 \text{s.t. } \quad & \sum_{i=1}^m c[i] \cdot X[i] \leq B \\
 & X_{\mathcal{P}}[j] \leq \sum_{i|e_i \in p_j} X[i], \forall j \\
 & 0 \leq X[i] \leq 1, \forall i \\
 & 0 \leq X_{\mathcal{P}}[j] \leq 1, \forall j
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 \text{LP10: } \min \quad & \sum_{j=1}^k \alpha[j] + B \cdot \gamma + \sum_{i=1}^m \delta[i] \\
 \text{s.t. } \quad & \alpha[j] + \beta[j] \geq f[j], \forall j \\
 & \delta[i] + c[i] \cdot \gamma \geq \sum_{j|p_j \ni e_i} \beta[j], \forall i
 \end{aligned} \tag{10}$$

The optimum solution of **LP9**, **LP10** must satisfy the following complementary slackness conditions:

$$\left\{ \begin{array}{ll} (X_{\mathcal{P}}[j] - 1)\alpha[j] = 0, \forall j & \dots(i) \\ (X_{\mathcal{P}}[j] - \sum_{i|e_i \in p_j} X[i])\beta[j] = 0, \forall j & \dots(ii) \\ (\sum_{i=1}^m c[i] \cdot X[i] - B)\gamma = 0 & \dots(iii) \\ (X[i] - 1)\delta[i] = 0, \forall i & \dots(iv) \\ (\alpha[j] + \beta[j] - f[j])X_{\mathcal{P}}[j] = 0, \forall j & \dots(v) \\ (\delta[i] + c[i] \cdot \gamma - \sum_{j|p_j \ni e_i} \beta[j])X[i] = 0, \forall i & \dots(vi) \end{array} \right.$$

We interpret γ as $f_e[i]/c[i]$, i.e. the bang-per-buck of all the edges attacked. Condition (vi) implies that the adversary will only attack edges with the highest γ . Let C_{min} be the min-cost $s - t$ cut and c^* be its cost. It's natural to attack C_{min} , as every possible $s - t$ path goes through at least 1 edge in C_{min} . Since the flow amount on C_{min} is at least 1, the highest γ is at least $1/c^*$. In case there are multiple min-cuts, the adversary may choose any one of them to achieve the same attack level on any possible paths, and different choices of min-cut will result in different optimal adversary strategies and different Nash equilibria, though the final utility will be the same. In the following example where the numbers denote the cost of edges, attacking edges in C_{min} yields the highest $\gamma = 1/4$.



Without loss of generality, we assume $B < c^*$, for otherwise the attack may simply cover every edge in C_{min} . It is possible to attack edges in C_{min} to obtain the attack on paths $X_{\mathcal{P}}$ that forces a rational designer to design the flow satisfying the following **Property A1**:

Property A1 : $\begin{cases} X_{\mathcal{P}}[j] = \chi, \forall j \text{ s.t. } f[j] > 0 \\ X_{\mathcal{P}}[j] \geq \chi, \forall j \end{cases}$, where χ is a constant $\in [0, 1]$

With this attack pattern, the designer does not need to deviate from her current routing.

Lemma 12. *Let f be the designer's flow strategy. Given an attack X , that defines $X_{\mathcal{P}}$, and which satisfies **Property A1** w.r.t f , the designer's flow strategy f is stable w.r.t. X , i.e. the designer has no incentive to change the flow design f .*

Proof. The designer’s current loss is χ since every path that carries flow is attacked by χ . Given any other flow pattern f' , the loss is $\sum_j X_{\mathcal{P}}[j]f'[j] \geq \sum_j \chi \cdot f'[j] = \chi$.

In other words, the current flow design uses only paths from the designer’s support set based on the attack, no other flow pattern will reduce the loss.

The following lemma describes a flow design that satisfies specific properties on C_{min} .

Lemma 13. *Given any graph $G(V, E)$, the designer is able to route a flow pattern f^* in $O(nm)$ steps, satisfying the following **Property A2**, where c^* is the cost of C_{min} .*

$$\textbf{Property A2: } \begin{cases} \forall e_i \in E, & f_e^*[i] \leq c[i]/c^* \\ \forall e_i \in C_{min}, & f_e^*[i] = c[i]/c^* \end{cases}$$

Proof. We create a new flow network as follows. Use the same vertex set V and for each undirected edge $e = (u, v) \in E$, create directed edges (u, v) and (v, u) , both with capacity c_e . Now we compute a maximum $s - t$ flow denoted by f^* in $O(nm)$ steps [13]. By the max-flow min-cut theorem, the flow amount is c^* since C_{min} is the min-cost $s - t$ cut. This flow satisfies the following:

$$\begin{cases} \forall e_i \in E, & f_e^*[i] \leq c_e \\ \forall e_i \in C_{min}, & f_e^*[i] = c_e \end{cases}$$

Now after scaling the flow amount on each edge by $1/c^*$, we get an $s - t$ flow amounting to 1, satisfying **Property A2**. Note that such flow design is not unique, so technically the optimal designer strategy and the Nash Equilibrium is not unique, but they all contribute to the same final utility. In case there are multiple min-cuts, the way of flow construction also makes sure of the optimality no matter which min-cut the adversary chooses to attack.

We observe that there is no edge in C_{min} carrying any flow from T to S , since f^* is a scaled max flow. We decompose the flow pattern into a set of paths \mathcal{P}' , $|\mathcal{P}'| \leq m$, such that $\forall p_j \in \mathcal{P}', f^*[j] > 0$. The decomposition can be computed in $O(nm)$ steps. We ignore any path $p_j \notin \mathcal{P}'$ as it carries no flow, therefore won’t be considered by the adversary. This allows the adversary to focus on a reduced set of paths \mathcal{P}' that doesn’t have exponential size.

Given the cut C_{min} , define an attack X_{min} of full coverage on C_{min} to be

$$\begin{cases} X_{min}[i] = 1, \forall i|e_i \in C_{min} \\ X_{min}[i] = 0, \forall i|e_i \notin C_{min} \end{cases}$$

That is, every edge in the cut is fully attacked. Note that since $B < c^*$ this attack is not within the budget.

The following lemma describes the attack at Nash equilibrium.

Lemma 14. *Let f^* be a flow design satisfying **Property A2**. Then the best response of the adversary is as follows*

$$\begin{cases} X^* = (B/c^*) \cdot X_{min} \\ X_{\mathcal{P}'}^*[j] = B/c^*, \forall j|p_j \in \mathcal{P}' \end{cases}$$

Proof. The response specified in the lemma only attacks edges from C_{min} with the same level. $X_{\mathcal{P}'}^*$ is calculated correctly because every path in \mathcal{P}' goes through only one edge in C_{min} . We show the optimality of the attack via complementary slackness.

We provide the following solution:

$$\begin{cases} \delta[i] = 0, \forall e_i \\ \alpha[j] = 0, \forall p_j \\ \beta[j] = f^*[j], \forall p_j \\ \gamma = 1/c^* \end{cases}$$

Complementary slackness conditions (i)(ii)(iv)(v) are clearly satisfied.

Condition (iii) is satisfied since $\sum_i X^*[i]c[i] = \sum_{i|e_i \in C_{min}} X^*[i]c[i] = \frac{B}{c^*} \cdot \sum_{i|e_i \in C_{min}} c[i] = \frac{B}{c^*} \cdot c^* = B$.

Condition (vi) is satisfied since

1. $\forall i|e_i \notin C_{min}, X^*[i] = 0$.
2. $\forall i|e_i \in C_{min}, \delta[i] + c[i] \cdot \gamma - \sum_{j|p_j \ni e_i} \beta[j] = c[i] \cdot \gamma - \sum_{j|p_j \ni e_i} f[j] = c[i] \cdot \gamma - f_{e_i} = \frac{c[i]}{c^*} - \frac{c[i]}{c^*} = 0$.

Therefore given the flow pattern, the proposed attack is optimum for the adversary.

Theorem 2. *f^* , the design from **Lemma 13**, and X^* , the attack from **Lemma 14**, form a Nash equilibrium for the **Network Routing Adversarial Game**. The Nash equilibrium can be computed in $O(nm)$ steps.*

5 r-Commodity Network Routing Adversarial Game

In this section we consider the game where there are r designers and 1 adversary. Recall the problem definition from the section on notations. The adversary attacks edges to capture part of the flow from all designers in the similar fashion.

Given the flow pattern f from the designers, the adversary solves the following program with its dual:

$$\begin{aligned}
 \text{LP11: } \max \quad & \sum_{j=1}^k f[j] \cdot X_{\mathcal{P}}[j] & \text{LP12: } \min \quad & \sum_{j=1}^k \alpha[j] + B \cdot \gamma + \sum_{i=1}^m \delta[i] \\
 \text{s.t. } \quad & \sum_{i=1}^m c[i] \cdot X[i] \leq B & \text{s.t. } \quad & \alpha[j] + \beta[j] \geq f[j], \forall j \\
 & X_{\mathcal{P}}[j] \leq \sum_{i|e_i \in \mathcal{P}_j} X[i], \forall j & & \delta[i] + c[i] \cdot \gamma \geq \sum_{j|p_j \ni e_i} \beta[j], \forall i \\
 & 0 \leq X \leq 1, \forall i & & \\
 & 0 \leq X_{\mathcal{P}} \leq 1, \forall j & &
 \end{aligned} \tag{11}$$

The optimum solution of **LP11**, **LP12** must satisfy the following complementary slackness conditions:

$$\left\{ \begin{array}{ll}
 (X_{\mathcal{P}}[j] - 1)\alpha[j] = 0, \forall j & \dots(i) \\
 (X_{\mathcal{P}}[j] - \sum_{i|e_i \in \mathcal{P}_j} X[i])\beta[j] = 0, \forall j & \dots(ii) \\
 (\sum_{i=1}^m c[i]X[i] - B)\gamma = 0 & \dots(iii) \\
 (X[i] - 1)\delta[i] = 0, \forall i & \dots(iv) \\
 (\alpha[j] + \beta[j] - f[j])X_{\mathcal{P}}[j] = 0, \forall j & \dots(v) \\
 (\delta[i] + c[i]\gamma - \sum_{j|p_j \ni e_i} \beta[j])X[i] = 0, \forall i & \dots(vi)
 \end{array} \right.$$

While we can solve the case of r -commodities using linear programming, the time complexity is high. We focus on the case when $r = 2$, where combinatorial algorithms exist. By using the costs as capacities of the edges and route max-flow with corresponding source and sink, we can identify min-cost cuts C_1, C_2 , and C_{12} which is a set of edges with minimum cost whose removal disconnects all $s_1 - t_1$ and $s_2 - t_2$ paths. Let c_1, c_2, c_{12} denote the costs of the cuts respectively. We ignore the case where the min-cut is not unique, as we discussed that such case doesn't affect the final utility in the previous section. We define attacks with full coverage of these cuts X_1, X_2, X_{12} in the same way in the previous section. We state below Hu's **Max Bi-flows Min-cut Theorem** [9] which shows feasibility of routing in the following subsections.

Theorem 3. *It is feasible to route a 2-commodity flow with requirement (r_1, r_2) iff*

$$\left\{ \begin{array}{l}
 r_1 \leq c_1 \\
 r_2 \leq c_2 \\
 r_1 + r_2 \leq c_{12}
 \end{array} \right.$$

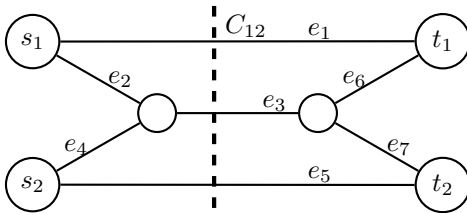
Without loss of generality, we assume $c_1 \leq c_2$. Since $c_{12} \leq c_1 + c_2$, we have $\frac{c_{12}}{2} \leq c_2$. We further assume $B \leq c_{12}$, for otherwise the attack may simply use X_{12} as her attack to capture the entire flow. Since the min-cost cuts carry different costs, the adversary attempts to capture the most beneficial one of the

them then turn to the rest. Based on the relationship between c_1 and c_{12} we divide the problem into several subcases.

Example: We use the same flow network with different cost on edges for the illustrative example in each following subsection. For simplicity we assume the cost of e_6, e_7 is $+\infty$; they will never be attacked. Note that in this network the direction of flow is obvious. The paths are $\mathcal{P} = \{p_1, p_2, p_3, p_4\}$, where $p_1 = (e_1)$, $p_2 = (e_2 \rightarrow e_3 \rightarrow e_6)$, $p_3 = (e_4 \rightarrow e_3 \rightarrow e_7)$, $p_4 = (e_5)$. The Nash equilibria are computed by linear programming.

Case(I) $\frac{c_{12}}{2} \leq c_1 \leq c_2$

In this case attacking C_{12} is more beneficial than attacking C_1 or C_2 . Following is an example of equilibrium where $B = 5, c_1 = 5, c_2 = 5, c_{12} = 8$. At the equilibrium, the adversary only attacks $C_{12} = \{e_1, e_3, e_5\}$ which covers boths designers' flow as C_{12} has better bang-per-buck than C_1 and C_2 .



edge	e_1	e_2	e_3	e_4	e_5
$c[i]$	2	3	4	3	2
$f_e[i]$	1/2	1/2	1	1/2	1/2
$X[i]$	5/8	0	5/8	0	5/8
path	p_1	p_2	p_3	p_4	
$f[j]$	1/2	1/2	1/2	1/2	
$X_{\mathcal{P}}[j]$	5/8	5/8	5/8	5/8	

The following lemma describes a 2-commodity flow design that satisfies specific properties on C_{12} .

Lemma 15. *In Case(I) the designers are able to route a 2-commodity flow pattern f^* with requirement (1, 1) in $O(n^3)$ steps, satisfying the following **PROPERTY A3***

$$\text{Property A3: } \begin{cases} \forall e_i \in E, & f_e^*[i] \leq 2c[i]/c_{12} \\ \forall e_i \in C_{12}, & f_e^*[i] = 2c[i]/c_{12} \end{cases}$$

We decompose the flow pattern into a set of paths $\mathcal{P}'_1, \mathcal{P}'_2$ with $|\mathcal{P}'_1| \leq m, |\mathcal{P}'_2| \leq m$. Let $\mathcal{P}' = \mathcal{P}'_1 \cup \mathcal{P}'_2$. The decomposition can be computed in $O(nm)$ steps. The adversary will focus the attack on \mathcal{P}' .

The following lemma describes the attack at Nash equilibrium.

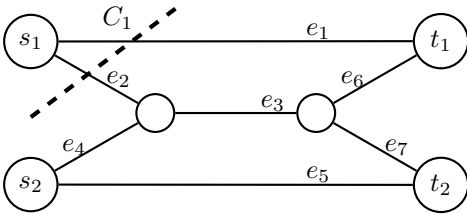
Lemma 16. *Let f^* be a flow design satisfying **PROPERTY A3**, the best response of the adversary is as follows:*

$$\begin{cases} X^* = (B/c_{12}) \cdot X_{12} \\ X_{\mathcal{P}}^*[j] = B/c_{12}, \forall j|p_j \in \mathcal{P}' \end{cases}$$

Lemma 17. When $\frac{c_{12}}{2} \leq c_1 \leq c_2$, f^* and attack X^* , defined in **Lemma 16**, describe a Nash equilibrium solution for the 2-Designer Network Routing Game computable in $O(n^3)$ steps.

Case(II) $c_1 < \frac{c_{12}}{2} \leq c_2$ and $B \leq c_1$

In this case the adversary attacks only C_1 since there is not enough budget to cover C_1 entirely and move onto C_{12} . Following is an example of equilibrium where $B = 4, c_1 = 5, c_2 = 8, c_{12} = 11$. At the equilibrium, the adversary only attacks $C_1 = \{e_1, e_2\}$ which only covers Designer 1’s flow as C_1 has better bang-per-buck than C_{12} and C_2 .



edge	e_1	e_2	e_3	e_4	e_5
$c[i]$	2	3	4	3	5
$f_e[i]$	1/2	1/2	2/3	1/6	5/6
$X[i]$	4/5	4/5	0	0	0
path	p_1	p_2	p_3	p_4	
$f[j]$	1/2	1/2	1/6	5/6	
$X_{\mathcal{P}}[j]$	4/5	4/5	0	0	

The following lemma describes the 2-commodity flow design that satisfies specific properties on C_1 .

Lemma 18. In **Case(II)** the designers are able to route a 2-commodity flow pattern f^* with requirement (1,1) in $O(n^3)$ steps, satisfying the following **PROPERTY A4**

$$\text{Property A4} : \begin{cases} \forall e_i \in E, & f_e^*[i] \leq c[i]/c_1 \\ \forall e_i \in C_1, & f_e^*[i] = c[i]/c_1 \end{cases}$$

We decompose the flow pattern into a set of paths $\mathcal{P}'_1, \mathcal{P}'_2$ with $|\mathcal{P}'_1| \leq m, |\mathcal{P}'_2| \leq m$. Let $\mathcal{P}' = \mathcal{P}'_1 \cup \mathcal{P}'_2$. The decomposition can be computed in $O(nm)$ steps. The adversary will focus the attack on \mathcal{P}' .

The following lemma describes the attack at Nash equilibrium.

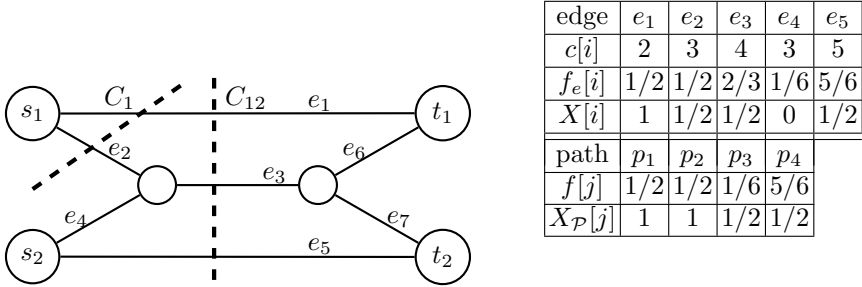
Lemma 19. Let f^* be a flow design satisfying **PROPERTY A4**, the best response of the adversary is as follows

$$\begin{cases} X^* = (B/c_1) \cdot X_1 \\ X_{\mathcal{P}'}^*[j] = B/c_1, \forall j|p_j \in \mathcal{P}'_1 \\ X_{\mathcal{P}'}^*[j] = 0, \forall j|p_j \in \mathcal{P}'_2 \end{cases}$$

Lemma 20. When $(c_1 < \frac{c_{12}}{2} \leq c_2) \& (B \leq c_1)$, f^* and attack X^* , defined in **Lemma 19**, provide a Nash equilibrium solution for the 2-Designer Network Routing Game computable in $O(n^3)$ steps.

Case(III) $c_1 < \frac{c_{12}}{2} \leq c_2$ and $c_1 < B \leq c_{12}$

In this case the adversary’s attack is a combination of X_1 and X_{12} to have C_1 fully attacked and C_{12} partially attacked. Following is an example of equilibrium where $B = 7, c_1 = 5, c_2 = 8, c_{12} = 11$. Note that the attack comprises of $C_1 = \{e_1, e_2\}$ and $C_{12} = \{e_1, e_3, e_5\}$. After capturing the most beneficial cut C_1 , the adversary has budget left so she moves on the attack C_{12} while maintaining the attack on C_1 to still be full. Therefore the adversary’s benefit is 1.5 which includes 1 from Designer 1 and 0.5 from Designer 2.



The flow design f^* in this case is decomposed into 2 flow routing components, f^1 and f^2 , i.e. $f^* = f^1 + f^2$. The following lemma describes f^2 .

Lemma 21. *In Case(III) the designers are able to route a 2-commodity flow pattern f^2 with requirement $(\frac{c_1}{c_{12}-c_1}, 1)$ in $O(n^3)$ steps, satisfying the following **PROPERTY A5***

$$\text{Property A5: } \begin{cases} \forall e_i \in E, & f_e^2[i] \leq c[i]/(c_{12} - c_1) \\ \forall e_i \in C_1 \cup C_{12}, & f_e^2[i] = c[i]/(c_{12} - c_1) \end{cases}$$

Proof. Use the costs as capacities of edges. Since $c_1 \leq \frac{c_{12}}{2}$, by **Theorem 3** it is possible to compute a feasible 2-commodity flow f^2 with requirement $(c_1, c_{12}-c_1)$ in $O(n^3)$ steps. This flow satisfies the following:

$$\begin{cases} \forall e_i \in E, & f_e^2[i] \leq c[i] \\ \forall e_i \in C_1 \cup C_{12}, & f_e^2[i] = c[i] \end{cases}$$

Scale f^* down by $\frac{1}{c_{12}-c_1}$, and it satisfies **PROPERTY A5** with requirement $(\frac{c_1}{c_{12}-c_1}, 1)$.

We decompose the flow pattern into sets of paths $\mathcal{P}'_1, \mathcal{P}'_2$ for Designer 1 and 2 respectively, with $|\mathcal{P}'_1| \leq m, |\mathcal{P}'_2| \leq m$. Let $\mathcal{P}' = \mathcal{P}'_1 \cup \mathcal{P}'_2$. The decomposition can be computed in $O(nm)$ steps. The adversary will focus the attack on \mathcal{P}' .

Lemma 22. *No path in \mathcal{P}'_2 goes through C_1 .*

Proof. Consider the max 2-commodity flow f^2 we obtained in **Lemma 21**, before scaling. C_1 is a min cut for commodity 1 with capacity c_1 , while the total flow amount of commodity 1 is also c_1 . Therefore $\forall e \in C_1, f_e^2$ contains only commodity 1. Since \mathcal{P}'_2 contains paths with positive flow of commodity 2, no path in \mathcal{P}'_2 goes through C_1 , i.e. $\forall p \in \mathcal{P}'_2, p \cap C_1 = \emptyset$.

Let p_l be one arbitrary path in \mathcal{P}'_1 , we define flow f^1 as follows:

$$f^1 = \begin{cases} f^1[l] = \frac{c_{12} - 2c_1}{c_{12} - c_1} \\ f^1[j] = 0, \forall j \neq l \end{cases}$$

Note that f^1 contains only commodity 1, its flow amount is $(\frac{c_{12}-2c_1}{c_{12}-c_1}, 0)$.

Let $f^* = f^1 + f^2$ be a 2-commodity flow, it satisfies the requirement $(\frac{c_{12}-2c_1}{c_{12}-c_1}, 0) + (\frac{c_1}{c_{12}-c_1}, 1) = (1, 1)$. Let $\theta = \frac{B-c_1}{c_{12}-c_1}$.

Lemma 23. *Given the flow design f^* , the best response of the adversary is as follows*

$$\begin{cases} X^* = (1 - \theta) \cdot X_1 + \theta \cdot X_{12} \\ X_{\mathcal{P}'_1}^*[j] = 1, \forall j | p_j \in \mathcal{P}'_1 \\ X_{\mathcal{P}'_2}^*[j] = \theta, \forall j | p_j \in \mathcal{P}'_2 \end{cases}$$

Proof. We construct the following dual variables of **LP12**

$$\begin{cases} \delta[i] = 0, \forall e_i \\ \alpha[j] = f^1[j], \forall j \\ \beta[j] = f^2[j], \forall j \\ \gamma = \frac{1}{c_{12} - c_1} \end{cases}$$

We proceed to show that all complementary slackness conditions of **LP11**, **LP12** are satisfied.

Condition (i) is satisfied since

1. $\forall p_j \in \mathcal{P}'_1, X_{\mathcal{P}'_1}[j] = 1$.
2. $\forall p_j \in \mathcal{P}'_2, \alpha[j] = 0$.

Condition (ii) is satisfied since

1. $\forall p_j \in \mathcal{P}'_1,$

$$\begin{aligned} X_{\mathcal{P}'_1}[j] - \sum_{i|e_i \in p_j} X[i] &= 1 - \sum_{i|e_i \in p_j} X[i] = 1 - (\sum_{i|e_i \in (p_j \cap C_1)} X[i] + \sum_{i|e_i \in (p_j \cap C_{12})} X[i]) \\ &= 1 - ((1 - \theta) + \theta) = 0 \end{aligned}$$

2. $\forall p_j \in \mathcal{P}'_2,$

$$\begin{aligned} X_{\mathcal{P}'_2}[j] - \sum_{i|e_i \in p_j} X[i] &= \theta - \sum_{i|e_i \in p_j} X[i] = \theta - (\sum_{i|e_i \in (p_j \cap C_1)} X[i] + \sum_{i|e_i \in (p_j \cap C_{12})} X[i]) \\ &= \theta - (\sum_{i|e_i \in \emptyset} X[i] + \sum_{i|e_i \in (p_j \cap C_{12})} X[i]) = \theta - (0 + \theta) = 0 \end{aligned}$$

Condition (iii) is satisfied since

$$\begin{aligned} \sum_i X[i] - B &= \left(\sum_{i|e_i \in X_1} c[i]X[i] + \sum_{i|e_i \in X_{12}} c[i]X[i] \right) - B = ((1-\theta) \cdot \sum_{i|e_i \in X_1} c[i] + \theta \cdot \sum_{i|e_i \in X_{12}} c[i]) - B \\ &= ((1-\theta) \cdot c_1 + \theta \cdot c_{12}) - B = c_1 + \theta(c_{12} - c_1) - B = 0 \end{aligned}$$

Condition (iv) is satisfied since $\delta[i] = 0, \forall i$.

Condition (v) is satisfied since $\alpha[j] + \beta[j] - f^*[j] = f^1[j] + f^2[j] - (f^1[j] + f^2[j]) = 0, \forall j$.

Condition (vi) is satisfied since

1. $\forall e_i \notin (C_1 \cup C_{12}), X[i] = 0$.
2. $\forall e_i \in (C_1 \cup C_{12}), c[i]\gamma - \sum_{j|p_j \ni e_i} \beta[j] = c[i]\gamma - f_e^2 = c[i]\gamma - \frac{c[i]}{c_{12}-c_1} = c[i] \cdot \frac{1}{c_{12}-c_1} - \frac{c[i]}{c_{12}-c_1} = 0$

Since all the conditions are satisfied, **LP11**, **LP12** are at optimum with the given variables. The adversary is not incentivized to deviate.

Lemma 24. *When $(c_1 < \frac{c_{12}}{2} \leq c_2) \& (c_1 < B \leq c_{12})$, f^* and attack X^* , defined in **Lemma 23**, describe a Nash equilibrium solution for the 2-Designer Network Routing Game that can be computed in $O(n^3)$ steps.*

Proof. Since C_1 and C_{12} are cuts of commodity 1, every possible path of Designer 1 goes through both C_1 and C_{12} at least once, thus $\forall p_j \in \mathcal{P}_1, X_{\mathcal{P}}[j] \geq (1-\theta) + \theta = 1$. By the decomposition of f^2 and construction of $f^1, \forall p_j \in \mathcal{P}_1 | f[j] > 0, X_{\mathcal{P}}[j] = (1-\theta) + \theta = 1$. Therefore for Designer 1, the attack satisfies **PROPERTY A1** with $\chi = 1$.

Every possible path of Designer 2 goes through C_{12} at least once, therefore the attack satisfies

PROPERTY A1 with $\chi = \theta$ for Designer 2.

By **Lemma 12**, neither of the designers wants to move from the current flow design.

Since neither the designers nor the adversary wants to move from the current strategy, this is a Nash equilibrium.

Combining the results of **Lemmas 17, 20, 24**, we come to the following conclusion:

Theorem 4. *For any input of the 2-Commodity Network Routing Adversarial Game, there exists a Nash equilibrium computable in $O(n^3)$ steps.*

6 Conclusions

In this paper we have illustrated how to determine Nash equilibrium for adversarial routing problems with player-specific costs on the edges. For the case when adversaries and designers are budget constrained and the graph is bipartite, we provide an efficient solution. We also provide efficient polynomial solutions to the general graph network routing adversary games, when the costs on the resources are not player-specific, and when there are 1 and 2 designers routing flow. It would be of interest to extend this to the more general case. This work was partially supported by NSF grant No: 2028274.

A Appendix

A.1 Proofs

Proof of Lemma 1

Proof.

$$\begin{aligned} B_D[k] - B_D[k-1] &= \frac{CD(k)}{C(k)} - \frac{CD(k-1)}{C(k-1)} = \frac{C(k-1) \cdot CD(k) - C(k) \cdot CD(k-1)}{C(k) \cdot C(k-1)} \\ &= \frac{c[k]}{C(k) \cdot C(k-1)} \cdot (d[k] \cdot C(k-1) - CD(k-1)) \end{aligned}$$

$$d[k] > d[i], \forall i < k \implies (d[k] \cdot C(k-1) - CD(k-1)) > 0 \implies B_D[k] - B_D[k-1] > 0$$

Proof of Lemma 2

Proof. Denote the attacks used in calculating $B_A[t-1]$ and $B_A[t]$ by X' and X'' respectively.

$$\begin{cases} X'[i] = \frac{d[t-1] - d[i]}{d[t-1] - d[1]}, & \forall i = 1 \dots t-1 \\ X''[i] = \frac{d[t] - d[i]}{d[t] - d[1]}, & \forall i = 1 \dots t \end{cases} \implies$$

$$\begin{aligned} X''[i] - X'[i] &= \frac{d[t]d[i] + d[t-1]d[1] - d[t-1]d[i] - d[t]d[1]}{(d[t] - d[1])(d[t-1] - d[1])} \\ &= \frac{(d[i] - d[1])(d[t] - d[t-1])}{(d[t] - d[1])(d[t-1] - d[1])} > 0, \forall i = 2 \dots t-1 \end{aligned}$$

Therefore

$$\begin{cases} X'[1] = X''[1] = 1 \\ X'[t] = X''[t] = 0 \\ X'[i] < X''[i], \\ \forall i = 2 \dots t-1 \end{cases} \implies B_A[t-1] = \sum_{i=1}^m c[i]X'[i] < \sum_{i=1}^m c[i]X''[i] = B_A[t]$$

Proof of Lemma 3

Proof.

1. $\gamma \geq 0$

The numerator:

$$B_D[\hat{k}] = \frac{CD(\hat{k})}{C(\hat{k})} \leq \frac{d[\hat{k}]C(\hat{k})}{C(\hat{k})} = d[\hat{k}], d[\hat{k}] \geq B_D[\hat{k}] \geq B_D \implies d[\hat{k}] - B_D \geq 0$$

The denominator:

$$d[\hat{k}] \geq d[i], \forall i < \hat{k} \implies d[\hat{k}]C(k-1) - CD(\hat{k}-1) > d[\hat{k}]C(k-1) - d[\hat{k}]C(k-1) = 0$$

2. $0 \leq f^*[i] \leq 1, \forall i$
 $\forall i = 1 \dots \hat{k} - 1, f^*[i] \geq 0$, since $f^*[i] = c[i]\gamma$.
 For $f^*[\hat{k}]$, the numerator:
 $B_D > B_D[\hat{k} - 1] \implies B_D \cdot C(\hat{k} - 1) > \frac{CD(\hat{k}-1)}{C(\hat{k}-1)}C(\hat{k} - 1) \implies B_D \cdot C(\hat{k} - 1) - CD(\hat{k} - 1) > 0$.
 The denominator > 0 is proven above.
 Since $f^*[i] \geq 0, \sum_{i=1}^{\hat{k}} f^*[i] = 1, f^*[i] \leq 1, \forall i$. Therefore $0 \leq f^*[i] \leq 1, \forall i$.
3. $f^*[i] \leq c[i]\gamma, \forall i$
 $\forall i = 1 \dots \hat{k} - 1, f^*[i] = c[i]\gamma$.
 $c[\hat{k}]\gamma - f^*[\hat{k}] = \frac{c[\hat{k}](d[\hat{k}] - B_D) - B_D \cdot C(\hat{k}-1) + CD(\hat{k}-1)}{c[\hat{k}]S(\hat{k}-1) - SQ(\hat{k}-1)}$
 The numerator:
 $c[\hat{k}](d[\hat{k}] - B_D) - B_D \cdot C(\hat{k} - 1) + CD(\hat{k} - 1) = CD(\hat{k}) - B_D \cdot C(\hat{k}) \geq CD(\hat{k}) - B_D[\hat{k}] \cdot C(\hat{k}) = CD(\hat{k}) - \frac{CD(\hat{k})}{C(\hat{k})}C(\hat{k}) = 0$.
 The denominator > 0 is proven above.
 Thus $\forall i = 1 \dots \hat{k}, f^*[i] \leq c[i]\gamma$.

Therefore f^*, γ are feasible.

Proof of Lemma 4

Proof.

1. $X_D, \Omega \geq 0$
 The numerators > 0 obviously.
 The denominator:
 $d[\hat{k}] > d[i], \forall i < k \implies d[\hat{k}]C(\hat{k}) - CD(\hat{k}) > d[\hat{k}]C(\hat{k}) - d[\hat{k}]C(\hat{k}) = 0$
2. $0 \leq X^*[i] \leq 1$
 The numerator:
 $d[\hat{k}] > d[i], \forall i < k \implies (d[\hat{k}] - d[i])B_A > 0$
 The denominator > 0 is proven above. Therefore $X^*[i] \geq 0$.
 Since $X^*[1] \geq X^*[i], \forall i = 1 \dots \hat{k}$, we need only to show $X^*[1] \leq 1$.

$$\begin{aligned} X^*[1] &= \frac{d[\hat{k}] - d[1]}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})} B_A \leq \frac{d[\hat{k}] - d[1]}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})} B_A[\hat{k}] \\ &= \frac{d[\hat{k}] - d[1]}{d[\hat{k}]C(\hat{k}) - CD(\hat{k})} \cdot \frac{d[\hat{k}]C(\hat{k}) - CD(\hat{k})}{d[\hat{k}] - dc[1]} = 1 \end{aligned}$$

Thus $0 \leq X^*[i] \leq 1, \forall i = 1 \dots \hat{k}$.

Therefore X^*, X_D, Ω are feasible.

Proof of Lemma 5

Proof. The complementary slackness conditions (ii)(iii)(v) are obviously satisfied due to the equations.

$\forall i = 1 \dots \hat{k} - 1, c[i]\gamma = f[i], \forall i = \hat{k} \dots m, X^*[i] = 0$. Therefore condition (i) is satisfied.

$\forall i = 1 \dots \hat{k}, X_D - X^*[i] = d[i]\Omega, \forall i = \hat{k} + 1 \dots m, f^*[i] = 0$. Therefore condition (iv) is satisfied.

$\forall i = 1 \dots m, \delta[i] = 0$. Therefore condition (vi) is satisfied.

All the complementary slackness conditions are met, the proposed solution is the optimal.

Proof of Lemma 6

Proof.

1. $\gamma \geq 0$

The numerator: $B_D - d[1] \geq 0$ by assumption.

The denominator: $CD(t) - d[1]C(t) = \sum_{i=1}^t c[i](d[i] - d[1]) > 0$.

2. $\delta[1] \geq 0$

The denominator > 0 is proven above.

Recall that \hat{k} is the largest integer with $B_D(\hat{k} - 1) < B_D$. Since $t \geq \hat{k}$, the numerator $CD(t) - B_D C(t) \geq CD(t) - B_D[t]C(t) = CD(t) - \frac{CD(t)}{C(t)}C(t) = 0$.

3. $0 \leq f^* \leq 1$

Since $\gamma, \delta[1] \geq 0, f^* \geq 0$. Since $\sum_{i=1}^t f^*[i] = 1, f^* \leq 1$.

Therefore $f^*, \gamma, \delta[1]$ are feasible.

Proof of Lemma 7

Proof.

1. $\Omega \geq 0$

The denominator: $c[i] > c[1], \forall i > 1 \implies CD(t) - d[1]C(t) > 0$.

The numerator: $C(t) - B_A > C(t) - B_A[t + 1] = C(t) - \sum_{i=1}^t c[i]X[i] + 0 \cdot X[t + 1] > C(t) - \sum_{i=1}^t c[i] \cdot 1 = C(t) - C(t) = 0$.

2. $X_D \geq 0$

$X_D - X^*[1] = d[1]\Omega \implies X_D = 1 + d[1]\Omega \geq 0$.

3. $0 \leq X^*[i] \leq 1, \forall i = 1 \dots t$

Since $X_D - X^*[i] = d[i]\Omega, X^*[1] > X^*[2] > \dots > X^*[t]. X^*[1] = 1 \implies X^*[i] \leq 1$.

To show $X^*[i] \geq 0$, it suffices to show $X^*[t] \geq 0$.

The denominator > 0 is proven above.

The numerator: $CD(t) + (d[t] - d[1])B_A - d[t]C(t) \geq CD(t) + (d[t] - d[1])B_A[t] - d[t]C(t) = CD(t) + (d[t] - d[1])\frac{d[t]C(t) - CD(t)}{d[t] - d[1]} - d[t]C(t) = 0$.

Thus $0 \leq X^*[i] \leq 1, \forall i = 1 \dots t$.

Therefore X_D, Ω, X^* are feasible.

Proof of Lemma 8

Proof. The complementary slackness conditions (ii)(iii)(v) are obviously satisfied due to the equations.

$\forall i = 1 \dots t, c[i]\gamma + \delta[i] = f^*[i], \forall i = t+1 \dots m, X^*[i] = 0$. Therefore condition (i) is satisfied.

$\forall i = 1 \dots t, X_D - X^*[i] = d[i]\Omega, \forall i = t+1 \dots m, f^*[i] = 0$. Therefore condition (iv) is satisfied.

$X^*[1] = 1, \forall i = 2 \dots m, \delta[i] = 0$. Therefore condition (vi) is satisfied.

All the complementary slackness conditions are met, the proposed solution is the optimal.

Proof of Lemma 9

Proof.

1. $\gamma \geq 0$

The numerator: $B_D - d[1] \geq 0$ by assumption.

The denominator: $CD(m) - d[1]C(m) = \sum_{i=1}^m c[i](d[i] - d[1]) > 0$.

2. $\delta[1] \geq 0$

The denominator > 0 is proven above.

The numerator:

Recall that \hat{k} is the largest integer with $B_D(\hat{k} - 1) < B_D$. Since $m \geq \hat{k}$, the numerator $CD(m) - B_D \cdot C(m) \geq CD(m) - B_D[m]C(m) = CD(m) - \frac{CD(m)}{C(m)}C(m) = 0$.

3. $0 \leq f^* \leq 1$

Since $\gamma, \delta[1] \geq 0, f^* \geq 0$. Since $\sum_{i=1}^t f^*[i] = 1, f^* \leq 1$.

Therefore $f^*, \gamma, \delta[1]$ are feasible.

Proof of Lemma 10

Proof.

1. $\Omega \geq 0$

The numerator $S(m) - B_A \geq 0$ by assumption.

The denominator $CD(m) - d[1]C(m) = \sum_{i=1}^m c[i](d[i] - d[1]) > 0$.

2. $X_D \geq 0$

$X_D - X^*[1] = c[1]\Omega \implies X_D = 1 + c[1]\Omega \geq 0$.

3. $0 \leq X^*[i] \leq 1, \forall i = 1 \dots m$

Since $X_D - X^*[i] = d[i]\Omega, X^*[1] > X^*[2] > \dots > X^*[m]$. $X^*[1] = 1 \implies X^*[i] \leq 1$.

To show $X^*[i] \geq 0$, it suffices to show $X^*[m] \geq 0$.

The denominator > 0 is proven above.

The numerator:

$$\begin{aligned}
 & CD(m) + (d[m] - d[1])B_A - d[m]C(m) \geq CD(m) + (d[m] - d[1])B_A[m] - d[m]C(m) \\
 & = CD(m) + (d[m] - d[1])\frac{d[m]C(m) - CD(m)}{d[m] - d[1]} - d[m]C(m) = 0
 \end{aligned}$$

Thus $0 \leq X^*[i] \leq 1, \forall i = 1 \dots m$.

Therefore X_D, Ω, X^* are feasible.

Proof of Lemma 11

Proof. The complementary slackness conditions $(ii)(iii)(v)$ are obviously satisfied due to the equations.

$\forall i = 1 \dots m, c[i]\gamma + \delta[i] = f^*[i]$. Therefore condition (i) is satisfied.

$\forall i = 1 \dots m, X_D - X^*[i] = d[i]\Omega$. Therefore condition (iv) is satisfied.

$X^*[1] = 1, \forall i = 2 \dots m, \delta[i] = 0$. Therefore condition (vi) is satisfied.

All the complementary slackness conditions are met, the proposed solution is the optimal.

Proof of Theorem 2

Proof. Since every possible path in the network has to go through C_{min} at least once, $X_{\mathcal{P}}[j] \geq B/c^*, \forall j$. The attack above satisfies **Property A1** with $\chi = B/c^*$, therefore the adversary’s gain is B/c^* , and the designer doesn’t want to change the flow design and use any path $p_j \notin \mathcal{P}'$.

Since neither the designer nor the adversary wants to move from the current strategy, this is a Nash equilibrium.

Proof of Lemma 15

Proof. Use the costs as capacities of edges. Since $\frac{c_{12}}{2} \leq c_1 \leq c_2$, by **Theorem 3** it is possible to compute a feasible 2-commodity flow f^* with requirements $(\frac{c_{12}}{2}, \frac{c_{12}}{2})$ in $O(n^3)$ steps [10]. This flow satisfies the following:

$$\begin{cases} \forall e_i \in E, & f_e^*[i] \leq c[i] \\ \forall e_i \in C_{12}, & f_e^*[i] = c[i] \end{cases}$$

Scale f^* down by $\frac{2}{c_{12}}$, and it satisfies **PROPERTY A3** and each designer’s flow sums up to be 1.

Proof of Lemma 16

Proof. We construct the following dual variables feasible for **LP12**, satisfying complementary slackness conditions.

$$\begin{cases} \delta[i] = 0, \forall e_i \\ \alpha[j] = 0, \forall p_j \\ \beta[j] = f^*[j], \forall p_j \\ \gamma = 2/c_{12} \end{cases}$$

Since the adversary’s program does not distinguish the 2 designers, by substituting C^* in the proof by C_{12} , the proof becomes identical to that of **Lemma 14**.

Proof of Lemma 17

Proof. Every possible path of each designer goes through C_{12} at least once, therefore this attack satisfies **PROPERTY A1** with $\chi = \frac{B}{c_{12}}$. By **Lemma 12** the designers don’t want to change the flow designs. The adversary’s gain is $\frac{2B}{c_{12}}$.

Since neither the designers nor the adversary wants to move from the current strategy, this is a Nash equilibrium.

Proof of Lemma 18

Proof. Use the costs as capacities of edges. Since $c_1 \leq \frac{c_{12}}{2}$, by **Theorem 3** it is possible to compute a feasible 2-commodity flow f^* with requirements (c_1, c_1) in $O(n^3)$ steps. This flow satisfies the following:

$$\begin{cases} \forall e_i \in E, & f_e^*[i] \leq c[i] \\ \forall e_i \in C_1, & f_e^*[i] = c[i] \end{cases}$$

Scale f^* down by $\frac{1}{c_1}$, and it satisfies **PROPERTY A4** and each designer’s flow sums up to be 1.

Proof of Lemma 19

Proof. We construct the following dual variables of **LP12**

$$\begin{cases} \delta[i] = 0, \forall e_i \\ \alpha[j] = 0, \forall p_j \\ \beta[j] = f^*[j], \forall p_j \\ \gamma = 1/c_1 \end{cases}$$

We proceed to show that all complementary slackness conditions of **LP11**, **LP12** are satisfied.

Condition (i)(iv)(v) are clearly satisfied.

Condition (ii) is satisfied since

1. $\forall p_j \in \mathcal{P}'_1$, since every path goes through C_1 only once, $\sum_{i|e_i \in p_j} X[i] = B/c_1 = X_{\mathcal{P}}[j]$.
2. $\forall p_j \in \mathcal{P}'_2$, since no edge or path of Designer 2 is attacked, $\sum_{i|e_i \in p_j} X[i] = X_{\mathcal{P}}[j] = 0$.

Condition (iii) is satisfied since

$$\sum_{i=1}^m c[i]X[i] = \sum_{i|e_i \in C_1} c[i]X[i] = B/c_1 \cdot \sum_{i|e_i \in C_1} c[i] = B/c_1 \cdot c_1 = B$$

Condition (vi) is satisfied since

1. $\forall e_i \notin C_1, X[i] = 0$.
2. $\forall e_i \in C_1,$

$$\delta[i] + c[i]\gamma - \sum_{j|p_j \ni e_i} \beta[j] = c[i]\gamma - \sum_{j|p_j \ni e_i} \beta[j] = c[i]\gamma - f_e^* = c[i]\gamma - \frac{c[i]}{c_1} = 0$$

Since all the conditions are satisfied, **LP11**, **LP12** are at optimum with the given variables. The adversary is not incentivized to deviate.

Proof of Lemma 20

Proof. No path in \mathcal{P}_2 is attacked, so Designer 2 doesn't want to deviate from the current flow design. Every possible path of Designer 1 goes through C_1 at least once, therefore the attack satisfies **PROPERTY A1** with $\chi = \frac{B}{c_1}$ for Designer 1.

The adversary's gain is $\frac{B}{c_1}$, and the designers don't want to change the flow designs.

Since neither the designers nor the adversary wants to change from the current strategy, this is a Nash equilibrium.

References

1. Ahmadinejad, A., Dehghani, S., Hajiaghayi, M., Lucier, B., Mahini, H., Seddighin, S.: From duels to battlefields: computing equilibria of blotto and other games. *Math. Oper. Res.* **44**, 1304–1325 (2019)
2. Barahona, F.: Packing spanning trees. *Math. Oper. Res.* **20**(1), 104–115 (1995)
3. Calinescu, G., Kapoor, S., Quinn, M., Shin, J.: Adversary games in secure/reliable network routing. In: Jain, R., Kannan, R. (eds.) *GameNets 2011*. LNCS, vol. 75, pp. 249–264. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30373-9_18
4. Chang, S.L., Piraveenan, M., Pattison, P., Prokopenko, M.: Game theoretic modelling of infectious disease dynamics and intervention methods: a review. *J. Biol. Dyn.* **14**(1), 57–89 (2020)
5. Chen, X., Deng, X., Teng, S.H.: Settling the complexity of computing two-player Nash equilibria. *J. ACM (JACM)* **56**(3), 14 (2009)
6. Grötschel, M., Lovász, L., Schrijver, A.: The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1**(2), 169–197 (1981)

7. Guan, P., He, M., Zhuang, J., Hora, S.C.: Modeling a multitarget attacker-defender game with budget constraints. *Decis. Anal.* **14**(2), 87–107 (2017)
8. Gueye, A., Walrand, J.C., Anantharam, V.: A network topology design game: How to choose communication links in an adversarial environment. In: Proc. of the 2nd International ICST Conference on Game Theory for Networks, GameNets, vol. 11, p. 5 (2011)
9. Hu, T.C.: Multi-commodity network flows. *Oper. Res.* **11**(3), 344–360 (1963)
10. Itai, A.: Two-commodity flow. *J. ACM (JACM)* **25**(4), 596–611 (1978)
11. Jin, K., Liu, M.: Multi-planner intervention in network games with community structures. arXiv preprint [arXiv:2103.16351](https://arxiv.org/abs/2103.16351) (2021)
12. Lye, K.W., Wing, J.M.: Game strategies in network security. *Int. J. Inf. Secur.* **4**(1–2), 71–86 (2005)
13. Orlin, J.B.: Max flows in $o(nm)$ time, or better. In: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, pp. 765–774. ACM (2013)
14. Papadimitriou, C.H.: On the complexity of the parity argument and other inefficient proofs of existence. *J. Comput. Syst. Sci.* **48**(3), 498–532 (1994)
15. Sandler, T.: Terrorism & game theory. *Simul. Gaming* **34**(3), 319–337 (2003)
16. Srivastava, V., et al.: Using game theory to analyze wireless ad hoc networks. *IEEE Commun. Surv. Tutorials* **7**(4), 46–56 (2005)
17. Wang, S., Liu, F., Shroff, N.: Non-additive security games. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
18. Wang, S., Shroff, N.: Security game with non-additive utilities and multiple attacker resources. *Proc. ACM Meas. Anal. Comput. Syst.* **1**(1), 1–32 (2017)
19. Washburn, A., Wood, K.: Two-person zero-sum games for network interdiction. *Oper. Res.* **43**(2), 243–251 (1995)
20. Xu, H.: The mysteries of security games: Equilibrium computation becomes combinatorial algorithm design. In: Proceedings of the 2016 ACM Conference on Economics and Computation, pp. 497–514 (2016)

Author Index

- Alam, Md Ibrahim Ibne 251
Alizadeh, Mahnoosh 310
Altman, Eitan 230
Anshevich, Elliot 251
- Bao, La 30
Brown, Philip N. 328
- Cai, Donghong 215
Cai, Yueping 83
Capone, Antonio 278
Chandan, Rahul 310
Chen, Xue 157
Chen, Yutian 132, 291
- Dai, Minghui 187
Dong, Zhicheng 44, 215
- Gao, Lynn 291
Garnaev, Andrey 55, 237
Guo, Shaoyong 100, 118
- Hota, Mohit 200
Huang, Jing 3
- Jiang, Tianyang 157
- Kapoor, Sanjiv 200, 339
Kar, Koushik 251
Kovenock, Dan 310
- Li, Jingwei 16
Li, Jingwen 150
Li, Tianchi 83
Li, Xin 73
Li, Yi 118
Liang, Wei 150
Liang, Yan 150
Lieto, Alessandro 278
Lin, Bin 187
Liu, Kang 157
Liu, Shaojun 16
Liu, Xu 157
Liu, Yan 30, 150
Long, Keping 175
- Lou, Zheng 16
Lu, Pengcheng 30, 100
Luo, Zhishen 187
Ly, Jennifer 132
- Malanchini, Ilaria 278
Mandelli, Silvio 278
Marden, Jason R. 310
- Ou, Qinghai 30
- Paarporn, Keith 310
Pan, Zhen 157
- Qian, Liping 187
Qiao, Chenbin 3
- Shang, Fangjian 73
Shao, Congzhang 100
Shao, Sujie 16, 30, 100, 118
Shende, Abhishek 263
- Tang, Bin 132, 291
Tong, Yuanzheng 3
Trappe, Wade 55, 237
- Vasal, Deepanshu 263
Vishwanath, Sriram 263
- Wang, Qianjun 118
Wang, Shuai 215
Wang, Tianshun 187
Wang, Wei 3
Wen, Zihao 44
Wu, Yuan 187
- Xiong, Ao 3
Xu, Yanqing 215
- Yao, Yanli 73
- Zhang, Guanghui 157
Zhang, Haijun 175
Zhang, Li 200
Zhang, Linlin 16
Zhang, Ningchi 30
Zhang, Yaomin 175

Zhang, Yi 339
Zhao, Jing 44
Zhao, Jun 150

Zheng, Tianren 73
Zhong, Cheng 30, 100, 118
Zhou, Weixi 44, 215