A MONETARY MECHANISM FOR STABILIZING COOPERATIVE DATA EXCHANGE WITH SELFISH USERS

A Thesis

by

ISHAN TYAGI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Alex Sprintson
Co-Chair of Committee,	Srinivas Shakkottai
Committee Member,	Gautam Natarajan
Head of Department,	Miroslav Begovic

December 2018

Major Subject: Computer Engineering

ABSTRACT

In this research, we address the stability issues in Cooperative Data Exchange (CDE), one of the central problems in wireless network coding. We consider a setting in which the users are selfish, i.e., would like to maximize their own utility. More specifically, we consider a setting where each user has a subset of packets in the ground set X, and wants all other packets in X. The users can exchange data by broadcasting coded or uncoded packets over a lossless channel, and monetary transactions are allowed between any pair of users. We define the utility of each user as the sum of two sub-utility functions: (i) the difference between the total payment received by the user and the total transmission rate of the user, and (ii) the difference between the total number of required packets by the user and the total payment made by the user. A rate-vector and payment-matrix pair (r, p) is said to stabilize the grand coalition (i.e., the set of all users) if (r, p) is Pareto-optimal over all minor coalitions (i.e., all proper subsets of users who collectively know all packets in X). Our goal is to design algorithms that compute a stabilizing rate-payment pair with minimum total sum-rate and minimum total sum-payment for any given instance of the problem. In this work, we propose two algorithms that maximize the sum of utility of all users (over all solutions), and one of the algorithms also maximizes the minimum utility among all users (over all solutions). The second algorithm requires a broker, where each user has to trust the broker and use the broker to exchange payments, whereas in the first algorithm there is no such requirement. In the first algorithm, the users directly compensate user broadcasting the packet in that particular round. Our scheme minimizes the total number of transmitted packets, as well as the total amount of payments. We also perform an extensive simulation study to evaluate the performance of our scheme in practical setting.

DEDICATION

To my family and friends

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Professor Alex Sprintson, for all the support, inspiration and mentoring that he provided throughout the course of my graduate studies. I would like to thank my co-advisor, Dr. Srinivas Shakkottai for all the valuable guidance regarding my career, research and coursework. I greatly benefitted from the informative discussions with both of them. I am grateful to Dr. Natarajan Gautam for serving on my thesis committee.

I wish to express my sincere appreciation and thanks to Dr. Anoosheh Heidarzadeh for mentoring me throughout the course of this thesis, being patient and allowing me to learn gradually and steadily. All the useful discussions with him taught me about how to approach a research problem, fragment it into smaller sub-problems and then most importantly, present it in a succint, meaningful way.

A special thanks to all the current and past NAPA lab members for their support and the valuable inputs.

Finally, I would like to thank my family, my friends, in College Station and in India, for the constant encouragement and the unwavering belief.

This thesis, for all its worth, is dedicated to all of them.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of Professor Alex Sprintson [advisor] and Professor Srinivas Shakkottai [co-advisor] of the Department of Electrical and Computer Engineering and Professor Gautam Natarajan of the Department of Industrial and Systems Engineering.

All the work conducted for the thesis was completed by the student in conjunction with Visiting Assistant Professor Anoosheh Heidarzadeh of the Department of Electrical and Computer Engineering.

Funding Sources

This work was done independently without outside financial support.

NOMENCLATURE

CDE	Cooperative Data Exchange
P2P	Peer-to-Peer
WSDE	Weakly Secure Data Exchange
MAC	Medium Access Control
TCP	Transmission Control Protocol
IC	Index Coding
MDS	Maximum Distance Separable

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

1. INTRODUCTION

Wireless technology has become the most sought after medium of gaining access to network. The novel technique of network coding greatly improves the performance of wireless networks. The traditional methods entail employing coding at the source nodes or the links to protect against the erasures/losses or to compress the redundant information. The network is handed the task of transmitting the information provided by the source nodes without any modification. In contrast to the traditional approach, the network coding technique allows for the mixing of the information received from different source nodes by the intermediary network nodes.

Ahlswede et al. proposed the technique of network coding in their seminal work [1]. Their work detailed the advantages of employing network coding over the traditional network operation methodology. All the packets are treated as the symbols of a finite field, which are transmitted by the intermediary nodes in the form of a coded combination. In order to decode the information in these coded combinations or to generate new packets, the receiving nodes have to apply the field operations.

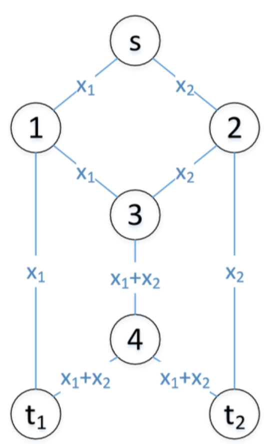

Figure 1.1: Network Coding Example

Figure 1.1 shows how network coding can be used to reduce the number of transmissions and henceforth, reduce the usage of network resources. This is an example showcasing multicast network coding. Each link in the network is capable of transmitting a single packet per channel use. Both the destination nodes, t_1 and t_2 need packets x_1 and x_2 from the source node s. As shown in the figure, only one channel use can suffice the needs of both the destination nodes, whereas, with the traditional routing approach, link (3,4) would have to be used twice, to transfer x_1 to t_2 and x_2 to t_1.

It was shown in [2] that employing network coding can enhance the robustness of the network in case of a link and node failure. After several intitial works on the wired networks, network coding found its application in wireless networks following the work done by Katti et al. [3]. They have introduced the *opportunistic listening* and *opportunistic coding* techniques. In opportunistic listening, wireless receivers store all the packets, for a short interval of time, transmitted over the wireless channel, regardless of packets' destination. Whereas, opportunistic coding is a strategy to improve network performance through coding by transmitting coded packets during the appropriate time slots. Following their seminal work, a lot of research has has been done in inspecting the advantages of network coding in wireless networks and the improvements in the network performance.

In this work, we shall go through the problem of Cooperative Data Exchange (CDE) in wireless networks with selfish users. El Rouayheb et al. [4] first proposed the CDE problem. It is a Peer-to-Peer (P2P) method of exchanging information, over a broadcast channel, among wireless clients. We consider the game-theoretic perspective of the CDE problem. We allow monetary transaction between the players. In our setting, each player has a utility function that captures the value fo the information gained through the exchange as well as the transmission cost. The problem is to find a *rate schedule*, i.e. the transmissions rate for each user, and a *payment schedule*, i.e. the payment made and received by each user during the transmission, to stabilize the *grand coalition* (i.e., set of all users). We propose two algorithms, each of which finds a solution that guarantees the stability of grand coalition for any instance. Both algorithms maximize the sum utility of all users, while one also maximizes the minimum utility among all users.

2. BACKGROUND AND RELATED WORK [1]

2.1 Background

Past few years have seen a lot of research being done in the field of cooperative communication in wireless networks. Cooperative communication can help us achieve space-time diversity, increased coverage, better data rates and energy efficiency. This kind of communication has the following advantages:

- Due to topological proximity of terminals, the transmissions within the close group of users is more reliable as compared to the transmission from the base station to any terminal,
- Local communication has a smaller footprint, hence allowing the user to utilize resources freely without interfering with the resources of the base station,
- Even if the connection with the base station is either too weak or unavailable after the initial phase of transmission, the users can communicate over the local network to recover the files.

The original setting of this problem considers a peer-to-peer data exchange scenario over a lossless broadcast channel. There is a group N of users and a ground set X of packets. Each user knows a subset of packets in X, and wants to learn the rest of packets in X. The users exchange their packets by broadcasting coded or uncoded versions of their packets, and the problem is to find a solution (i.e., the transmission rate of each user and the set of packets transmitted by each user) such that all users achieve omniscience with minimum total sum-rate. In this work, we revisit the CDE problem from a game-theoretic perspective where all users are selfish.

[1] *Reprinted with permission from "A Monetary Mechanism for Stabilizing Cooperative Data Exchange with Selfish Users" by A. Heidarzadeh, I. Tyagi, S. Shakkottai, and A. Sprintson, 2018, in *Proc. IEEE ISIT' 18, Jun. 2018 ©2018 IEEE.*

2.1.1 CDE with Selfish Clients

In this setting, there can be a monetary transaction between any pair of users, and the utility function of each user is defined as the sum of two sub-utility functions as follows: (i) the difference between the total payment the user receives from other users and its transmission rate, and (ii) the difference between the total number of packets the user wants and the total payment it makes to other users. Thinking of the sum of the transmission rate and the total payment being made by each user as its *cost* for participating in the exchange session, and thinking of the sum of the number of packets each user learns and the total payment being received by the user as its *gain* due to its participation in the exchange session, the utility function of each user is the surplus of the user.

The problem is to find a rate schedule $\{r_i\}_{i\in N}$ and a payment schedule $\{p_{i,j}\}_{i,j\in N}$ for the *grand coalition* (i.e., the set of all users) to achieve omniscience all together that is Pareto optimal, with respect to the utility function, over all *minor coalitions* (i.e., any proper subset of users who collectively know all packets in X). That is, a pair $(\{r_i\}_{i\in N}, \{p_{i,j}\}_{i,j\in N})$ is a *solution* if there is no pair $(\{\tilde{r}_i\}_{i\in S}, \{\tilde{p}_{i,j}\}_{i,j\in S})$ for any minor coalition S to achieve omniscience together such that the utility of some user(s) in S is strictly greater, and the utility of no user in S is less. Note that a solution *stabilizes* the grand coalition in that no minor coalition has incentive to break the grand coalition. The goal is to find a solution that minimizes the total sum-rate and the total sum-payment simultaneously.

In this work, we propose two algorithms, each of which finds a solution for any problem instance. Moreover, we show that both algorithms maximize the sum of utility of all users (over all solutions), and one of the algorithms also maximizes the minimum utility among all users (over all solutions).

2.2 Related Work

Since the seminal work by Ahlswede et al. [1], the field of network coding has much evolved. Starting work focused mainly on network code construction for multicast networks, modeled as graph with one or more source nodes and multiple sink nodes. The amount of work being done in wireless network applications of network coding is far less as compared to work done in wired networks. The most conventional ways of increasing throughput in wireless networks are mostly related to making the routing protocols more efficient or modifying the trannsport or MAC protocols [5–7]. Deb et al. [8] and Lun et al. [9] first explored the network coding applications in wireless network systems. However, both of the papers employed algorithms developed for wired multicast networks and hence, remained far from the problems specific to the wireless systems. Katti et al. [3], [10] made the first major breakthrough in the wireless network coding. They have proposed algorithms that leverage the broadcast nature of wireless networks. Specifically, they propose a framework in which wireless devices listen to packets sent by the neighboring devices, irrespective of the packets' destinations. This technique is called opportunistic listening. With the help of a coding scheme that can take advantage of this information, devices can improve the overall throughput by combining packets from various sources. The corresponding technique is called opportunistic coding. Both these techniques laid the foundation for network coding schemes in wireless domain, which could effectively utilize the broadcast properties of wireless channel. Authors in [11], propose different optimal opportunistic coding algorithms for various mesh networks. Settings with one-hop networks, consisting of only two nodes exchanging packets among themselves and one node capable of sending and receiving packets from both nodes, were studied in [12–15].

Another problem which received very wide attention from the researchers, first introduced in the context of satellite networks by Birk et al. in [16, 17], is Index Coding (IC). The clients in IC have some prior information through opportunistic learning and are in need of certain packets. The base station transmitting to the clients, can take advantage of the knowledge of the clients and in turn, can reduce the total transmissions required. After [17], researchers developed great insight into the properties of IC and followed with many papers. Bar-Yossef et al. [18] formulated IC as a graph theory problem. IC was eventually proved to be an NP-hard problem in [19] which is hard to approximate in [20].

Similar to IC, CDE is a setting of wireless network coding problem which employs opportunistic coding and listening. The side information present with the CDE clients is acquired through the technique of opportunistic learning. However, unlike IC, the clients in CDE perform opportunistic coding in order to get all of the required information with minimum transmission cost. After being first proposed by Rouayheb et al. [4], much work has been done with respect to the algorithms for this problem [21–24]. Reference [21] provided the randomized algorithm and [22] proposed the deterministic algorithm for CDE problem. Reference [23] gave a divide-and-conquer method to find the optimal coding strategy. A different version of the problem was given in [24], where fractional packets were considered and hence, submodularity property of the cut-set bounds was employed. Work in [25–27] examined cost, fairness and multi-hop network topology w.r.t. CDE. CDE is related to the secret key agreement problem, formulated in [28], in [29].

A coalition-game model for the CDE problem was recently proposed in [30]. This model differs from our work in two aspects: (i) the utility function is different from ours, and (ii) the criteria for the stability of the grand coalition is different from the Pareto optimality considered here.

Recently, in [31], there was some work being done on a related problem, where each user has two utility functions: its rate and its delay. They defined the stability of the grand coalition via the Pareto optimality with respect to both the rate and delay functions simultaneously. The researchers showed that over all the minor coalitions there does not exist any *non-monetary* mechanism (without the peer-to-peer payments) that stabilizes the grand coalition for all problem instances. This result is the motivation of this work on the design of a *monetary* mechanism for stabilizing the grand coalition for any problem instance.

3. STABILIZING COOPERATIVE DATA EXCHANGE PROBLEM [1]

3.1 Problem Setup

We consider the original setting of the cooperative data exchange (CDE) problem as follows. Consider a group of n users and a set of k packets $X \triangleq \{x_1, \dots, x_k\}$. Let $N \triangleq \{1, \dots, n\}$ and $K \triangleq \{1, \dots, k\}$. Initially, each user $i \in N$ has a subset X_i of the packets in X, and ultimately, the user i wants the rest of the packets $\overline{X}_i \triangleq X \setminus X_i$. The index set of packets in X_i for each user i is known by all other users. Also, without loss of generality, we assume that $X = \cup_{i \in N} X_i$. The objective of all users is to *achieve omniscience*, i.e., to learn all packets in X, via exchanging their packets by broadcasting (coded or uncoded) packets.

A subset S of users in N is a *coalition* if $\cup_{i \in S} X_i = X$. We refer to any coalition $S \subset N$ as a *minor coalition*, and refer to the coalition N as the *grand coalition*. Whenever we use the notation S for a subset of users, we assume that S is a coalition, unless explicitly noted otherwise.

Let $\mathbb{Z}_+$ be the set of non-negative integers. For any $S \subseteq N$, a rate vector $r \triangleq [r_1, \dots, r_n] \in \mathbb{Z}_+^n$ is *S-omniscience-achieving* if there exists a transmission scheme with each user $i \in S$ transmitting r_i (coded or uncoded) packets such that all users in S achieve omniscience, regardless of transmissions of the rest of the users. Note that, for any S-omniscience-achieving rate vector, random linear network coding (over a sufficiently large finite field) suffices as a transmission scheme for all users in S to achieve omniscience (with any arbitrarily high probability) [25].

For any $S \subseteq N$, we denote by $\mathcal{R}_S$ the set of all S-omniscience-achieving rate vectors r such that $r_i = 0$ for all $i \notin S$. For any arbitrary subset $S \subseteq N$ and any rate vector r, we define the sum-rate $r_S \triangleq \sum_{i \in S} r_i$ and $r_\emptyset \triangleq 0$. By a standard network coding argument [25], for any $S \subseteq N$, $r \in \mathcal{R}_S$ iff $r_{\tilde{S}} \geq |\cap_{j \in S \setminus \tilde{S}} \overline{X}_j|$, for every (non-empty) $\tilde{S} \subset S$.

[1] *Reprinted with permission from "A Monetary Mechanism for Stabilizing Cooperative Data Exchange with Selfish Users" by A. Heidarzadeh, I. Tyagi, S. Shakkottai, and A. Sprintson, 2018, in *Proc. IEEE ISIT' 18, Jun. 2018 ©2018 IEEE.*

We consider CDE under a monetary mechanism where there can be a payment from any user to any other user. For all $i, j \in N$, let $p_{i,j} \geq 0$ be the total payment from the user i to the user j, and let $p_{i,i} = 0$. For a payment matrix $p \triangleq [p_{i,j}]$, let $p_i^+ \triangleq \sum_{j \in N \setminus \{i\}} p_{j,i}$ and $p_i^- \triangleq \sum_{j \in N \setminus \{i\}} p_{i,j}$ be the total incoming payment of the user i and the total outgoing payment of the user i, respectively.

For any $S \subseteq N$, we denote by $\mathcal{P}_S$ the set of all payment matrices p such that $p_{i,j} = 0$ and $p_{j,i} = 0$ for all $i \in S$, $j \notin S$, i.e., there is no incoming payment to any user in S from any user out of S and there is no outgoing payment from any user in S to any user out of S. For any $S \subseteq N$, we define the sum-payment $p_S \triangleq \sum_{i,j \in S} p_{i,j}$. Note that $\sum_{i \in S} p_i^+ = \sum_{i \in S} p_i^- = p_S$ for all $p \in \mathcal{P}_S$.

Definition 1 (Utility). *For any $S \subseteq N$, any $r \in \mathcal{R}_S$, and any $p \in \mathcal{P}_S$, the utility of each user $i \in S$ is given by*

$$u_i(r, p) \triangleq (p_i^+ - r_i) + (|\overline{X}_i| - p_i^-),$$

where $u_i^+(r, p) \triangleq p_i^+ - r_i$ is the net utility due to the user i's contribution to the system, and $u_i^-(r, p) \triangleq |\overline{X}_i| - p_i^-$ is the net utility due to the system's contribution to the user i.

Note that the cost per transmission and the value per packet are assumed to be unity for all users.

The two functions $u_i^+(r, p)$ and $u_i^-(r, p)$ motivate the notion of rationality defined as follows.

Definition 2 (Rationality). *For any $S \subseteq N$, any $r \in \mathcal{R}_S$ and any $p \in \mathcal{P}_S$, the rate-payment pair (r, p) is* rational *if $u_i^+(r, p) \geq 0$ and $u_i^-(r, p) \geq 0$ for all $i \in S$.*

Hereafter, we focus on the rational rate-payment pairs only, and omit the term "rational" for brevity.

We assume that all the users are *selfish*, i.e., each user may or may not agree with its rate specified by a rate vector or its payments specified by a payment matrix. The goal is to find a rate-payment pair (r, p), $r \in \mathcal{R}_N$ and $p \in \mathcal{P}_N$, under which N is *stable*. We formally define the notion of stability based on the utility function as follows.

Definition 3 (Stability)**.** *For any rate-payment pair (r, p), $r \in \mathcal{R}_N$ and $p \in \mathcal{P}_N$, N is (r, p)-*stable *if there is not a rate-payment pair $(\tilde{r}, \tilde{p})$, $\tilde{r} \in \mathcal{R}_S$, and $\tilde{p} \in \mathcal{P}_S$, for some $S \subset N$, such that*

- $u_i(r, p) \leq u_i(\tilde{r}, \tilde{p})$ *for all $i \in S$, and*
- $u_i(r, p) < u_i(\tilde{r}, \tilde{p})$ *for some $i \in S$.*

The (r, p)-stability of the grand coalition is equivalent to the Pareto optimality of (r, p) over all minor coalitions.

Definition 4 (Feasibility)**.** *A rate-payment pair (r, p) is* feasible *if N is (r, p)-stable.*

Note that a feasible solution guarantees that no minor coalition of users has incentive to break the grand coalition.

Definition 5 (Optimality)**.** *A feasible (r, p) is* optimal *if there is not a feasible $(\tilde{r}, \tilde{p})$ such that $r_N > \tilde{r}_N$ or $p_N > \tilde{p}_N$.*

Note that, for an optimal solution, the sum-rate and the sum-payment are minimum among all feasible solutions.

The problem is to determine if an optimal solution exists for any given instance, and if so, to find such a solution.

3.2 Proposed Algorithms

Algorithm 1

In this section, we present an algorithm that, for any given instance, finds an optimal solution.

The algorithm begins with an all-zero rate vector $r = [r_i]_{i \in N}$ and an all-zero payment matrix $p = [p_{i,j}]_{i,j \in N}$, operates in rounds, and updates r and p over the rounds.

Algorithm 1: Algo1(n, k, $\{\mathrm{U}_i\}_{i=1}^n$, $\mathbb{F}_q$)

$N \leftarrow \{1,\dots,n\}$, $K \leftarrow \{1,\dots,k\}$
$r_i \leftarrow 0\ \forall i \in N$, $p_{i,j} \leftarrow 0\ \forall i,j \in N$
$l \leftarrow 1$, $\mathrm{V}_0 \leftarrow \emptyset$
while $\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1}) < k$ for some $i \in N$ **do**
    $T_l \leftarrow \{i \in N : \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1}) = \max_{i\in N} \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1})\}$
    Select an arbitrary user $t \in T_l$
    $R_l \leftarrow \{i \in N : \mathrm{U}_t \not\subseteq \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1})\}$
    Select an encoding vector $v_l \in \mathbb{F}_q^k$ such that $v_l^i = 0\ \forall \{i \in K : u_i \notin \mathrm{U}_t\}$ and $v_l \notin \mathrm{span}(\cup_{i\in R_l} \mathrm{U}_i \cup \mathrm{V}_{l-1})$
    Have the user t transmit the packet $y_l = \sum_{i\in K} v_l^i x_i$
    $r_t \leftarrow r_t + 1$
    $p_{i,t} \leftarrow p_{i,t} + 1/|R_l|\ \forall i \in R_l$
    $\mathrm{V}_l \leftarrow \mathrm{V}_{l-1} \cup v_l$
    $l \leftarrow l + 1$
end
return $r = [r_i]_{i\in N}$ and $p = [p_{i,j}]_{i,j\in N}$

For any (uncoded) packet x_i, $i \in K$, denote the (unit) encoding vector of x_i by $u_i \triangleq [u_i^1,\dots,u_i^k]$, where $u_i^i = 1$ and $u_i^j = 0$ for all $j \neq i$. For any (linearly coded) packet $y_j \triangleq \sum_{i\in K} v_j^i x_i$, where $v_j^i \in \mathbb{F}_q$ (for some finite field $\mathbb{F}_q$), denote the encoding vector of y_j by $v_j \triangleq [v_j^1,\dots,v_j^k]$.

Let U_i be the set of (unit) encoding vectors of packets in X_i, and V_l be the set of encoding vectors of all packets being transmitted by the end of the round l. Let $\mathrm{V}_0 \triangleq \emptyset$. We refer to $\mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_l)$ and $\dim(\mathrm{U}_i \cup \mathrm{V}_l)$ as the *knowledge* and the *size of knowledge* of the user i at the end of the round l, respectively, where $\mathrm{span}(\mathrm{V})$ and $\dim(\mathrm{V})$ denote the vector space of (linear) span (over $\mathbb{F}_q$) of a collection V of vectors in $\mathbb{F}_q^k$ and the dimension of $\mathrm{span}(\mathrm{V})$, respectively.

Consider an arbitrary round $l > 0$. Let T_l be the set of all users i with maximum $\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1})$. In the round l, the algorithm first selects an arbitrary user $t \in T_l$, and then the user t constructs (using its uncoded packets) and broadcasts a (coded) packet y_l (with encoding vector v_l).

Let R_l be the set of all users i such that $\mathrm{U}_t \not\subseteq \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1})$. The encoding vector

v_l of the packet y_l satisfies two conditions: (i) $v_l^i = 0\ \forall\{i \in K : u_i \notin \mathrm{U}_t\}$, and (ii) $v_l \notin \mathrm{span}(\cup_{i\in R_l}\mathrm{U}_i \cup \mathrm{V}_{l-1})$. (Such a vector $v_l \in \mathbb{F}_q^k$ always exists and it can be found in polynomial time using a randomized or a deterministic algorithm so long as $q \geq n \cdot k$ or $q \geq n$, respectively [21].) Note that R_l is the set of all users i whose knowledge at the beginning of the round l is not a superset of (initial) knowledge of the transmitting user t, and the encoding vector v_l of the packet y_l being transmitted by the user t in the round l is not known to any user $i \in R_l$ at the beginning of the round l. Thus, the transmission of the packet y_l increases the size of knowledge of any user $i \in R_l$ by one, and it does not change that of any user $i \notin R_l$.

Next, the algorithm increments r_t by 1 and increments $p_{i,t}$ by $1/|R_l|$ for all $i \in R_l$. At the end of the round l, the algorithm augments V_{l-1} by v_l, and constructs V_l, i.e., $\mathrm{V}_l = \mathrm{V}_{l-1} \cup \{v_l\}$. The rounds continue until the size of knowledge all users is k. Once the algorithm terminates, it returns the rate vector r and the payment matrix p.

Example below shows how this algorithm works for a group of users with side information. Consider a group of 4 users and a set of 6 packets, $\{x_1, \ldots, x_6\}$. Each figure below has the want-set, packets required by that user to gain access to the complete file, written inside it. For example, user 1 here has packet x_1 in its want-set or in other words, has access to packets in $X_1 = \{x_2, \ldots, x_6\}$. Each user i will start with its own subset X_i and wants the packets in $\overline{X_i}$. The index set of packets available to each user is known by the other users. Want-sets written in red depict the users who are still incomplete and those in black belong to the users who have gained access to all the packets/become complete by that moment.

Rate vector, payment matrix and utility vector belonging to this CDE are, $r = [2, 1, 0, 0]$, $p = \left[0, \frac{1}{3}, 0, 0; \frac{1}{3}, 0, 0, 0; \frac{5}{6}, \frac{1}{3}, 0, 0; \frac{5}{6}, \frac{1}{3}, 0, 0\right]$, $u(r,p) = [\frac{2}{3}, \frac{2}{3}, \frac{11}{6}, \frac{11}{6}]$, respectively.

Theorem 1. *The output of Algorithm 1 is optimal.*

Algorithm 2

In this section, we present an algorithm that for any given instance provides an optimal solution with maximum sum-utility and maximum min-utility among all optimal solutions.

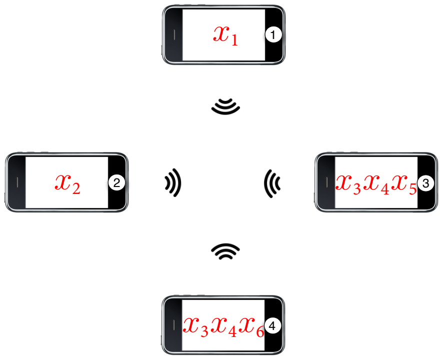

Figure 3.1: Users in the initial setting

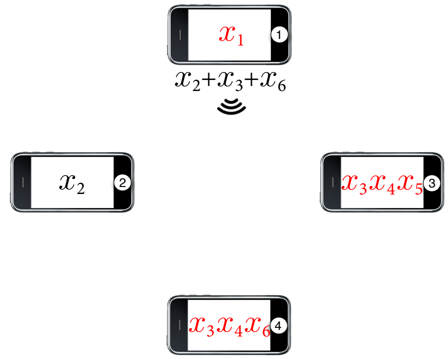

Figure 3.2: User 1 transmits in the first round

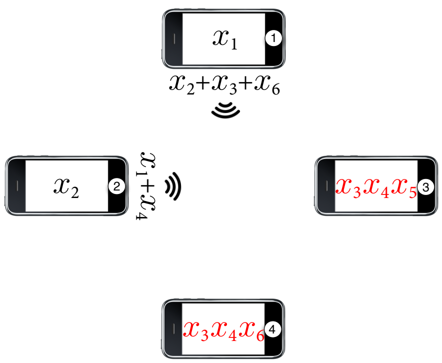

Figure 3.3: User 2 transmits in the second round

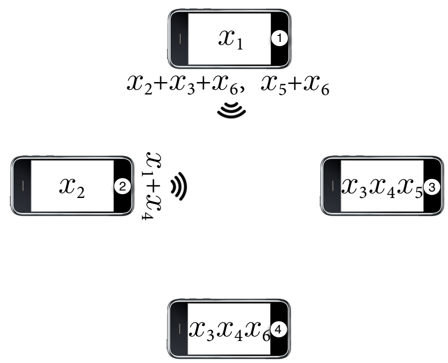

Figure 3.4: User 1 transmits again in the third round

Algorithm 2 is similar to Algorithm 1, and the only difference is in the set of users that make payments and the update rule of the payments in each round. We assume that there is a *broker* that collects the payment p_i^- by each user i, and returns the payment p_i^+ to each user i. The algorithm begins with all-zero payment vectors p^+ and p^-, and updates these vectors over the rounds as follows. Consider an arbitrary round $l > 0$. Let P_l be the set of users with maximum $|\overline{X}_i| - p_i^-$. Assuming that the user t transmits in the round l, the algorithm increments p_t^+ by 1 and increments p_i^- by $1/|P_l|$ for all $i \in P_l$.

In the example shown for algorithm 1, if we run algorithm 2, the rate vector, payment vectors and utility vector are, $r = [2, 1, 0, 0]$, $p^+ = [2, 1, 0, 0]$, $p^- = [0, 0, \frac{3}{2}, \frac{3}{2}]$, $u(r, p) = [1, 1, \frac{3}{2}, \frac{3}{2}]$, respectively.

Theorem 2. *The output of Algorithm 2 is optimal. Moreover, the output of Algorithm 2 has*

Algorithm 2: Algo2(n, k, $\{\mathrm{U}_i\}_{i=1}^{n}$, $\mathbb{F}_q$)

$N \leftarrow \{1, \ldots, n\}$, $K \leftarrow \{1, \ldots, k\}$
$r_i \leftarrow 0$, $p_i^+ \leftarrow 0$, $p_i^- \leftarrow 0\ \forall i \in N$
$l \leftarrow 1$, $\mathrm{V}_0 \leftarrow \emptyset$
while $\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1}) < k$ for some $i \in N$ **do**
  $T_l \leftarrow \{i \in N : \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1}) = \max_{i \in N} \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1})\}$
  $P_l \leftarrow \{i \in N : |\overline{X}_i| - p_i^- = \max_{i \in N}(|\overline{X}_i| - p_i^-)\}$
  Select an arbitrary user $t \in T_l$
  $R_l \leftarrow \{i \in N : \mathrm{U}_t \not\subseteq \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1})\}$
  Select an encoding vector $v_l \in \mathbb{F}_q^k$ such that $v_l^i = 0\ \forall \{i \in K : u_i \notin \mathrm{U}_t\}$ and $v_l \notin \mathrm{span}(\cup_{i \in R_l} \mathrm{U}_i \cup \mathrm{V}_{l-1})$
  Have the user t transmit the packet $y_l = \sum_{i \in K} v_l^i x_i$
  $r_t \leftarrow r_t + 1$
  $p_t^+ \leftarrow p_t^+ + 1$
  $p_i^- \leftarrow p_i^- + 1/|P_l|\ \forall i \in P_l$
  $\mathrm{V}_l \leftarrow \mathrm{V}_{l-1} \cup v_l$
  $l \leftarrow l + 1$
end
return $r = [r_i]_{i \in N}$ and $p = [p_i^+, p_i^-]_{i \in N}$

maximum sum-utility and maximum min-utility among all optimal solutions.

3.3 Proof of theorems

Proof of Theorem 1

In this section, we reserve the notations r and p for the outputs of Algorithm 1.

Lemma 1. *(r, p) is rational (i.e., $p_i^+ \geq r_i$ and $|\overline{X}_i| \geq p_i^-$ for all $i \in N$).*

Proof. By the procedure of Algorithm 1, $p_i^+ = r_i$ since the user i receives one unit of payment for each transmission it makes, and $|\overline{X}_i| \geq p_i^-$ since the user i pays at most one unit for each transmission that increases its size of knowledge, and it does not pay for any other transmission. □

Let N_s be the sth subset of users that achieve omniscience simultaneously, and let l_s be the round at which the users in N_s achieve omniscience. Note that the sets N_s are disjoint.

Denote by $N^{(s)}$ the set of all users in $N_1, \dots, N_s$. Let m be such that $N^{(m)} = N$. By using similar ideas as in the proof of [31, Lemma 4], the following result can be shown.

Lemma 2. *For any $s \in [m]$ and any $S \subseteq N^{(s)}$ such that $S \cap N_s \neq \emptyset$, we have $l_s \leq \tilde{r}_S$ for all $\tilde{r} \in \mathcal{R}_S$.*

Proof. Fix an arbitrary $s \in [m]$. Fix an arbitrary $S \subseteq N^{(s)}$ such that $S \cap N_s \neq \emptyset$, and an arbitrary $\tilde{r} \in \mathcal{R}_S$. Let $\{y_l\}_{1 \leq l \leq l_s}$ be the set of the algorithm's choice of packets being transmitted from the round 1 to the round l_s, and let $\{v_l\}_{1 \leq l \leq l_s}$ be the set of encoding vectors of these packets.

For any $S \subseteq N$, we say that a set of packets is S*-transmittable* if the encoding vector of each packet in the set lies in $\mathrm{span}(\mathrm{U}_i)$ for some $i \in S$. Let $\ell \triangleq \min(\tilde{r}_S, l_s)$. We prove by induction (on l) that, for every $1 \leq l \leq \ell$, there exists an S-transmittable set of $\tilde{r}_S - l + 1$ packets such that if they were transmitted after the transmission of all the packets in the set $\{y_1, \dots, y_{l-1}\}$, then S achieves omniscience.

For the base case of $l = 1$, there exists an S-transmittable set of $\tilde{r}_S$ packets such that if they were transmitted, then S achieves omniscience (since $\tilde{r} \in \mathcal{R}_S$). Next, consider an arbitrary round l, $1 < l \leq \ell$. Fix the set of packets $Y = \{y_1, \dots, y_{l-1}\}$. By the induction hypothesis, there exists an S-transmittable set of $\tilde{r}_S - l + 1$ packets such that if they were transmitted after the transmission of Y, then S achieves omniscience. Let $\tilde{Y} \triangleq \{\tilde{y}_l, \dots, \tilde{y}_{\tilde{r}_S}\}$ and $\tilde{V} \triangleq \{\tilde{v}_l, \dots, \tilde{v}_{\tilde{r}_S}\}$ be such a set of packets and the set of their encoding vectors, respectively. Assume that the algorithm selects the user t, which may or may not be in S, to transmit in the round l.

Since

$$\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1}) \geq k - \tilde{r}_S + l - 1$$

for all $i \in S$ (noting that, after the transmission of $Y \cup \tilde{Y}$, S achieves omniscience), and

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1}) \geq \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1})$$

for all $i \in N$ (noting that, in the round l, the size of the knowledge of the user t is greater than or equal to that of any other user $i \in N$), then

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1}) \geq k - \tilde{r}_S + l - 1.$$

If

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1}) = k - \tilde{r}_S + l - 1,$$

then the user t cannot transmit in the round l since the user t needs the set of all the packets in $\tilde{Y}$ so as to achieve omniscience. This is, however, a contradiction (by assumption). Thus,

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1}) > k - \tilde{r}_S + l - 1,$$

and consequently, $\tilde{Y}$ contains some packet $\tilde{y}$ such that its encoding vector $\tilde{v} \in \mathrm{span}(\mathrm{U}_t \cup \mathrm{V}_{l-1})$. Fix such a packet $\tilde{y}$ and its encoding vector $\tilde{v}$. Note that, after the transmission of $Y \cup \tilde{Y} \setminus \{\tilde{y}\}$, the user t achieves omniscience (i.e., $\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) = k$), and any user $i \in S$, $i \neq t$, needs no more than one packet so as to achieve omniscience (i.e., $\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) \geq k - 1$ for all $i \in S$, $i \neq t$). (The deletion of one packet decreases the size of knowledge of any user by at most one.)

Consider an arbitrary $i \in S, i \neq t$. We consider two cases: (i) $v_l \in \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\})$, and (ii) $v_l \notin \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\})$. In the case (i), since

$$v_l \in \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\})$$

and

$$v_l \in \mathrm{span}(\mathrm{U}_t \cup \mathrm{V}_{l-1}),$$

then

$$\mathrm{span}(\mathrm{U}_t \cup \mathrm{V}_{l-1}) \subseteq \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}),$$

or equivalently,

$$\mathrm{span}(\mathrm{U}_t \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) \subseteq \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}).$$

Thus,

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) \leq \dim(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}),$$

or equivalently,

$$\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) = k$$

since

$$\dim(\mathrm{U}_t \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) = k.$$

Thus, after the transmission of $Y \cup \tilde{Y} \setminus \{\tilde{y}\}$, the user i achieves omniscience. In the case (ii), since

$$v_l \notin \mathrm{span}(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\})$$

and

$$\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \tilde{V} \setminus \{\tilde{v}\}) \geq k-1,$$

then

$$\dim(\mathrm{U}_i \cup \mathrm{V}_{l-1} \cup \{v_l\} \cup \tilde{V} \setminus \{\tilde{v}\}) = k.$$

Thus, after the transmission of $Y \cup \{y_l\} \cup \tilde{Y} \setminus \{\tilde{y}\}$, the user i achieves omniscience. By (i) and (ii), it follows that S achieves omniscience after the transmission of $Y \cup \{y_l\} \cup \tilde{Y} \setminus \tilde{y}$. Thus, there exists an S-transmittable set of $\tilde{r}_S - l$ packets $\tilde{Y} \setminus \tilde{y}$ such that if they were transmitted after the transmission of $Y \cup y_l$, then S achieves omniscience. This completes the inductive proof.

From the above result, it follows that S achieves omniscience by the algorithm's choice of packets $\{y_l\}_{1 \leq l \leq \ell}$ being transmitted from the round 1 to the round ℓ. Now there are two cases: (i) $l_s > \tilde{r}_S$, and (ii) $l_s \leq \tilde{r}_S$. In the case (i), $\ell = \tilde{r}_S$, and hence, all users in S must achieve omniscience by the round ℓ ($= \tilde{r}_S$). This is, however, a contradiction since some user(s) in S, particularly any user in $S \cap N_s$, achieves omniscience in the round l_s ($> \ell$) (by definition). Note that $S \cap N_s \neq \emptyset$ (by assumption). In the case (ii), $\ell = l_s$, and the lemma follows directly. This completes the proof. $\square$

Lemma 3. *(r, p) is feasible (i.e., N is (r, p)-stable).*

Proof. The proof follows by contradiction. Suppose that (r, p) is not feasible (i.e., N is not (r, p)-stable). Thus, there exists $\tilde{r} \in \mathcal{R}_S$ and $\tilde{p} \in \mathcal{P}_S$ for some $S \subset N$ such that $u_i(r, p) \leq u_i(\tilde{r}, \tilde{p})$ for all $i \in S$, and $u_i(r, p) < u_i(\tilde{r}, \tilde{p})$ for some $i \in S$. Thus,

$$\sum_{i \in S} u_i(\tilde{r}, \tilde{p}) > \sum_{i \in S} u_i(r, p).$$

Note that

$$\sum_{i \in S} u_i(\tilde{r}, \tilde{p}) = \sum_{i \in S} p_i^+ - \sum_{i \in S} \tilde{r}_i + \sum_{i \in S} |\overline{X}_i| - \sum_{i \in S} p_i^-.$$

Since $\sum_{i \in S} p_i^+ = \sum_{i \in S} p_i^-$ for all $p \in \mathcal{P}_S$, then

$$\sum_{i \in S} u_i(\tilde{r}, \tilde{p}) = \sum_{i \in S} |\overline{X}_i| - \sum_{i \in S} \tilde{r}_i.$$

Since $r_i = p_i^+$ for all $i \in N$, then $\sum_{i \in S} r_i = \sum_{i \in S} p_i^+$. Thus,

$$\sum_{i \in S} u_i(r, p) = \sum_{i \in S} |\overline{X}_i| - \sum_{i \in S} p_i^-.$$

Putting these arguments together, we get

$$\sum_{i \in S} p_i^- > \sum_{i \in S} \tilde{r}_i. \tag{3.1}$$

Let $s \in [m]$ be such that $S \subseteq N^{(s)}$ and $S \cap N_s \neq \emptyset$. Note that all the users in S achieve omniscience by the round l_s. By the structure of the proposed algorithm, one unit of payment is made in each round (each user in R_l pays $1/|R_l|$ units of payment in the round l), and no user pays in any round after it achieves omniscience (if the user i is complete at the beginning of the round l, then $i \notin R_l$). Thus, it is easy to see that

$$\sum_{i\in S} p_i^- \leq l_s.$$

Moreover, by the result of Lemma 2, it follows that

$$l_s \leq \sum_{i\in S} \tilde{r}_i$$

for all $\tilde{r} \in \mathcal{R}_S$. By combining these two inequalities, we get

$$\sum_{i\in S} p_i^- \leq \sum_{i\in S} \tilde{r}_i. \tag{3.2}$$

By comparing (3.1) and (3.2), we arrive at a contradiction. Thus, N is (r, p)-stable, as was to be shown. □

Lemma 4 ([21]). *For any $\tilde{r} \in \mathcal{R}_N$, we have $\tilde{r}_N \geq r_N$.*

Proof. The proof can be found in [21]. □

Lemma 5. *(r, p) is optimal (i.e., there is not a feasible $(\tilde{r}, \tilde{p})$ such that $r_N > \tilde{r}_N$ or $p_N > \tilde{p}_N$).*

Proof. Consider an arbitrary feasible $(\tilde{r}, \tilde{p})$, $\tilde{r} \in \mathcal{R}_N$ and $\tilde{p} \in \mathcal{P}_N$. We shall show that $\tilde{r}_N \geq r_N$ and $\tilde{p}_N \geq p_N$. By Lemma 4, $\tilde{r}_N \geq r_N$ for all $\tilde{r} \in \mathcal{R}_N$. Since $(\tilde{r}, \tilde{p})$ is feasible, then $(\tilde{r}, \tilde{p})$ is rational. Thus, $\tilde{p}_i^+ \geq \tilde{r}_i$ for all $i \in N$, and consequently, $\tilde{p}_N \geq \tilde{r}_N$. Note that $p_N = r_N$ since $p_i^+ = r_i$. Thus, $\tilde{p}_N \geq \tilde{r}_N \geq r_N = p_N$. This completes the proof. □

Proof of Theorem 2

In this section, we reserve the notations r and p for the outputs of Algorithm 2.

Lemma 6. *(r, p) is rational.*

Proof. Let $r_i(l)$, $p_i^+(l)$, and $p_i^-(l)$ be r_i, p_i^+, and p_i^- at the end of the round $l-1$, respectively. Note that $r_i = r_i(l_m+1)$, $p_i^+ = p_i^+(l_m+1)$, and $p_i^- = p_i^-(l_m+1)$. We will show that $p_i^+(l) \geq r_i(l)$ and $|\overline{X}_i| \geq p_i^-(l)$ for all $i \in N$ and all $l \in [l_m+1]$. Fix an arbitrary $l \in [l_m+1]$. By the procedure of Algorithm 2, $p_i^+(l) = r_i(l)$, and particularly, $p_i^+ = r_i$. We next show that $|\overline{X}_i| \geq p_i^-(l)$. The proof follows by contradiction. Suppose that $|\overline{X}_i| < p_i^-(l)$ for some i. Note that

$$\max_{i \in N} |\overline{X}_i| - p_i^-(l) = k - \min_{i \in N}(|X_i| + p_i^-(l)).$$

Thus,

$$P_l = \{i \in N : |X_i| + p_i^-(l) = \min_{i \in N}(|X_i| + p_i^-(l))\}.$$

By the procedure of Algorithm 2, $|X_i| + p_i^-(l)$ are the same for all i such that $p_i^-(l) > 0$, and $|X_i| + p_i^-(l) = |X_i| \leq k$ for all i such that $p_i^-(l) = 0$. Since $|X_i| + p_i^-(l) > k$ for some i (by assumption), then $|X_i| + p_i^-(l) > k$ for all i, and consequently, $p_i^-(l) > 0$ for all i (since $|X_i| \leq k$ for all i). Since $p_i^-(l)$ is non-decreasing in l for all i, then $|X_i| + p_i^- > k$ for all i, or equivalently, $p_i^- > |\overline{X}_i|$ for all i. Thus,

$$\sum_{i \in N} p_i^- > \sum_{i \in N} |\overline{X}_i|,$$

and consequently,

$$r_N > \sum_{i \in N} |\overline{X}_i|$$

since

$$\sum_{i \in N} p_i^- = \sum_{i \in N} p_i^+ = r_N.$$

This is, however, a contradiction since

$$r_N \leq \min_{i \in N} |\overline{X}_i| + \max_{i \in N} |\overline{X}_i|$$

(by the result of [4, Lemma 3]), and consequently,

$$r_N \leq \sum_{i \in N} |\overline{X}_i|.$$

Thus, $|\overline{X}_i| \geq p_i^-(l)$ for all i and all l, and particularly, $|\overline{X}_i| \geq p_i^-$ for all i. This completes the proof. $\square$

Lemma 7. (r, p) *is feasible.*

Proof. Take an arbitrary S such that $\mathcal{R}_S \neq \emptyset$ (i.e., all users in S can achieve omniscience together). By the same argument as in the proof of Lemma 3, it suffices to show that

$$\sum_{i \in S} p_i^- \leq \tilde{r}_S$$

for all $\tilde{r} \in \mathcal{R}_S$. Run Algorithm 2 over the set S, and denote by $(\tilde{r}, \tilde{p})$ the output. Let $\tilde{Y} = \{\tilde{y}_1, \dots, \tilde{y}_{\tilde{r}_S}\}$ and $\tilde{V} = \{\tilde{v}_1, \dots, \tilde{v}_{\tilde{r}_S}\}$ be the set of all packets being transmitted from the round 1 to the round $\tilde{r}_S$ and their encoding vectors, respectively. Note that $\tilde{r}_S$ is the minimum sum-rate that all users in S can achieve omniscience (by Lemma 4). Assume, without loss of generality, that $|X_1| \leq |X_2| \leq \cdots \leq |X_n|$. Define $i^\star \triangleq \min_{i \in S} i$, and $S^\star \triangleq \{i^\star, \dots, n\}$. Since $S \subseteq S^\star$, then $\mathcal{R}_{S^\star} \neq \emptyset$ (i.e., all users in $S^\star$ can achieve omniscience together). Moreover, run Algorithm 2 over the set $S^\star$, and denote by $(r^\star, p^\star)$ the output. Note that $p^\star_{S^\star} = r^\star_{S^\star}$ (by the result of Lemma 6). Let $Y^\star = \{y^\star_1, \dots, y^\star_{r^\star_{S^\star}}\}$ and $V^\star = \{v^\star_1, \dots, v^\star_{r^\star_{S^\star}}\}$ be the set of all packets being transmitted from the round 1 to the round $r^\star_{S^\star}$ and their encoding vectors, respectively.

First, we show that

$$r^\star_{S^\star} \leq \tilde{r}_S.$$

To do so, it suffices to show that all users in $S^\star \setminus S$ achieve omniscience after the reception of all packets in $\tilde{Y}$. The proof follows by contradiction. Consider an arbitrary user $i \in S^\star \setminus S$. Suppose that the user i does not achieve omniscience after the reception of all packets in

$\tilde{Y}$, i.e., $\dim(\mathrm{U}_i \cup \tilde{V}) < k$. Since $\dim(\mathrm{U}_i) \geq \dim(\mathrm{U}_{i^\star})$ and $\dim(\mathrm{U}_{i^\star}) \geq k - \tilde{r}_S$, then $\dim(\mathrm{U}_i) \geq k - \tilde{r}_S$. Thus, there exists some round l such that the encoding vector $\tilde{v}_l$ of the packet $\tilde{y}_l$ being transmitted by some user $t \in S$ is in the knowledge set of the user i prior to the round l, i.e.,

$$\mathrm{span}(\mathrm{U}_t) \subseteq \mathrm{span}(\mathrm{U}_i \cup \{\tilde{v}_1, \dots, \tilde{v}_{l-1}\}),$$

and consequently,

$$\mathrm{span}(\mathrm{U}_t \cup \{\tilde{v}_1, \dots, \tilde{v}_{l-1}\}) \subseteq \mathrm{span}(\mathrm{U}_i \cup \{\tilde{v}_1, \dots, \tilde{v}_{l-1}\}).$$

Thus,

$$\mathrm{span}(\mathrm{U}_t \cup \tilde{V}) \subseteq \mathrm{span}(\mathrm{U}_i \cup \tilde{V}).$$

Since $\dim(\mathrm{U}_t \cup \tilde{V}) = k$ and $\dim(\mathrm{U}_i \cup \tilde{V}) \geq \dim(\mathrm{U}_t \cup \tilde{V})$, then $\dim(\mathrm{U}_i \cup \tilde{V}) = k$. This is, however, a contradiction since $\dim(\mathrm{U}_i \cup \tilde{V}) < k$ (by assumption). Thus, all users in $S^\star \setminus S$ achieve omniscience after the reception of all packets in $\tilde{Y}$, and so, $r^\star_{S^\star} \leq \tilde{r}_S$.

Next, we show that

$$\sum_{i \in S^\star} p_i^- \leq p^\star_{S^\star}.$$

If $S^\star = N$, then

$$\sum_{i \in S^\star} p_i^- = p_N = r_N = r^\star_N = p^\star_N = p^\star_{S^\star}.$$

Now assume that $S^\star \neq N$. If for some l, the packet $y^\star_l$ being transmitted by the user $t \in S^\star$ does not increase the size of knowledge of the user $i \in N \setminus S^\star$ such that $\dim(\mathrm{U}_i \cup V^\star) < k$, then

$$\mathrm{span}(\mathrm{U}_t) \subseteq \mathrm{span}(\mathrm{U}_i \cup \{v^\star_1, \dots, v^\star_{l-1}\}),$$

and consequently,

$$\mathrm{span}(\mathrm{U}_t \cup \{v^\star_1, \dots, v^\star_{l-1}\}) \subseteq \mathrm{span}(\mathrm{U}_i \cup \{v^\star_1, \dots, v^\star_{l-1}\}).$$

Thus,

$$\mathrm{span}(\mathrm{U}_t \cup V^\star) \subseteq \mathrm{span}(\mathrm{U}_i \cup V^\star).$$

Since $\dim(\mathrm{U}_t \cup V^\star) = k$ and $\dim(\mathrm{U}_i \cup V^\star) \geq \dim(\mathrm{U}_t \cup V^\star)$, then $\dim(\mathrm{U}_i \cup V^\star) = k$. This yields a contradiction since $\dim(\mathrm{U}_i \cup V^\star) < k$ (by assumption). Thus, the packet $y^\star_l$ (for any l) increases the size of knowledge of all users in $N \setminus S^\star$ that do not achieve omniscience after the reception of all packets $y^\star_1, \dots, y^\star_{r^\star_{S^\star}}$.

Since the size of knowledge of each user $i \in N \setminus S^\star$ after the reception of all packets in $Y^\star$ is $\min\{|X_i| + r^\star_{S^\star}, k\}$, then the user i needs $k - \min\{|X_i| + r^\star_{S^\star}, k\}$ $(\leq k - \min\{|X_1| + r^\star_{S^\star}, k\})$ more packets to achieve omniscience. Thus, if the users in $S^\star$ continue to make transmissions after they all achieve omniscience, all users in $N \setminus S^\star$ achieve omniscience after the reception of at most $k - \min\{|X_1| + r^\star_{S^\star}, k\}$ more packets. Thus, all users in N achieve omniscience with at most $r^\star_{S^\star} + k - \min\{|X_1| + r^\star_{S^\star}, k\}$ total transmissions. Since r_N is the minimum sum-rate for all users in N to achieve omniscience, then

$$r^\star_{S^\star} + k - \min\{|X_1| + r^\star_{S^\star}, k\} \geq r_N.$$

We consider two cases: (i) $|X_1| + r^\star_{S^\star} \geq k$, and (ii) $|X_1| + r^\star_{S^\star} < k$.

In the case (i), we have

$$r^\star_{S^\star} \geq r_N = p_N \geq \sum_{i \in S^\star} p^-_i.$$

In the case (ii), we have

$$r^\star_{S^\star} + k - |X_1| - r^\star_{S^\star} = k - |X_1| \geq r_N.$$

Since $r_N \geq k - |X_1|$ (otherwise, the user 1 cannot achieve omniscience), then $r_N = k - |X_1|$. Let $c \triangleq \min_{i \in N}(|X_i| + p^-_i)$. If $c < |X_{i^\star}|$, then

$$\sum_{i \in S^\star} p^-_i \leq p^\star_{S^\star}$$

since $\sum_{i\in S^\star} p_i^- = 0$. Now, assume that $c \geq |X_{i^\star}|$. Recall that $|X_1| \leq |X_2| \leq \cdots \leq |X_{i^\star}| \leq \cdots \leq |X_n|$ (by assumption). Thus, $c \geq |X_i|$ for all $i \in N \setminus S^\star$. Note that

$$\sum_{i\in S^\star} p_i^- = r_N - \sum_{i\in N\setminus S^\star} (c - |X_i|)$$

and

$$p^\star_{S^\star} = r^\star_{S^\star}.$$

We need to show that

$$\sum_{i\in S^\star} p_i^- \leq p^\star_{S^\star}.$$

Thus it suffices to show that

$$r_N - \sum_{i\in N\setminus S^\star} (c - |X_i|) \leq r^\star_{S^\star}.$$

The proof follows by contradiction. Suppose that

$$r_N - \sum_{i\in N\setminus S^\star} (c - |X_i|) > r^\star_{S^\star}.$$

Since $r_N = k - |X_1|$ and $r^\star_{S^\star} \geq k - |X_{i^\star}|$ (otherwise, the user $i^\star$ cannot achieve omniscience), then

$$k - |X_1| - \sum_{i\in N\setminus S^\star} (c - |X_i|) > r^\star_{S^\star} \geq k - |X_{i^\star}|,$$

and consequently,

$$|X_{i^\star}| > |X_1| + \sum_{i\in N\setminus S^\star} (c - |X_i|).$$

Since

$$\sum_{i\in N\setminus S^\star} (c - |X_i|) = (i^\star - 1)c - (|X_1| + \cdots + |X_{i^\star - 1}|),$$

then

$$|X_2| + \cdots + |X_{i^\star}| > (i^\star - 1)c.$$

This is, however, a contradiction since $c \geq |X_i|$ for all $i \in [i^\star]$ (by assumption), and so,

$$(i^\star - 1)c \geq |X_2| + \cdots + |X_{i^\star}|.$$

Thus,

$$\sum_{i \in S^\star} p_i^- \leq p^\star_{S^\star}.$$

Moreover,

$$\sum_{i \in S} p_i^- \leq \sum_{i \in S^\star} p_i^-$$

since $S \subseteq S^\star$ (by definition). By combining the above arguments, it then follows that

$$\sum_{i \in S} p_i^- \leq \tilde{r}_S,$$

as was to be shown. □

Lemma 8. (r, p) *is optimal.*

Proof. The proof follows from the same argument as in the proof of Lemma 5, and hence omitted to avoid repetition. □

Lemma 9. *For any optimal* $(\tilde{r}, \tilde{p})$*, we have* $\sum_{i \in N} u_i(r, p) = \sum_{i \in N} u_i(\tilde{r}, \tilde{p})$ *and* $\min_{i \in N} u_i(r, p) \geq \min_{i \in N} u_i(\tilde{r}, \tilde{p})$.

Proof. The proof of the first part (i.e., maximum sum-utility) is straightforward. Take an arbitrary optimal $(\tilde{r}, \tilde{p})$. Since $\tilde{r}_N = r_N$ and $p, \tilde{p} \in \mathcal{P}_N$, then

$$\sum_{i \in N} u_i(\tilde{r}, \tilde{p}) = \sum_{i \in N} |\overline{X}_i| - \tilde{r}_N = \sum_{i \in N} |\overline{X}_i| - r_N = \sum_{i \in N} u_i(r, p).$$

For the proof of the second part (i.e., maximum min-utility), we need to show that

$$\min_{i \in N} u_i(\tilde{r}, \tilde{p}) \leq \min_{i \in N} u_i(r, p)$$

for any optimal $(\tilde{r}, \tilde{p})$. Take an arbitrary optimal $(\tilde{r}, \tilde{p})$. Since $\tilde{p}_i^+ = \tilde{r}_i$ (otherwise, $\tilde{p}_N > r_N = p_N$ since $\tilde{p}_i^+ \geq \tilde{r}_i$ (by rationality of $(\tilde{r}, \tilde{p})$), and so, $(\tilde{r}, \tilde{p})$ cannot be optimal), then $u_i(\tilde{r}, \tilde{p}) = |\overline{X}_i| - \tilde{p}_i^-$. Note that $\tilde{p}_N = p_N$. Let $c \triangleq \min_{i \in N}(|X_i| + p_i^-)$. Note that $|\overline{X}_i| - p_i^- = k - c$ if $c \geq |X_i|$, and $|\overline{X}_i| - p_i^- = |\overline{X}_i| = k - |X_i|$ if $c < |X_i|$. Thus, $u_i(r, p) = k - \max\{c, |X_i|\}$ for all $i \in N$. Since $|X_n| \geq |X_i|$ for all $i \in N$ (by assumption), then it follows that

$$\min_{i \in N} u_i(r, p) = k - \max\{c, |X_n|\}.$$

We consider two cases: (i) $c < |X_n|$, and (ii) $c \geq |X_n|$.

In the case (i), $\min_{i \in N} u_i(r, p) = k - |X_n| = |\overline{X}_n|$. If $\tilde{p}_n^- > 0$, then

$$u_n(\tilde{r}, \tilde{p}) = |\overline{X}_n| - \tilde{p}_n^- < |\overline{X}_n| = \min_{i \in N} u_i(r, p).$$

If $\tilde{p}_n^- = 0$, then $u_n(\tilde{r}, \tilde{p}) = |\overline{X}_n|$, and consequently,

$$\min_{i \in N} u_i(\tilde{r}, \tilde{p}) \leq u_n(\tilde{r}, \tilde{p}) = \min_{i \in N} u_i(r, p) = k - c.$$

In the case (ii), $\min_{i \in N} u_i(r, p) = k - c$. Suppose that $\min_{i \in N} u_i(\tilde{r}, \tilde{p}) > \min_{i \in N} u_i(r, p)$. Let $j \in N$ be such that $|\overline{X}_j| - \tilde{p}_j^- = \min_{i \in N} u_i(\tilde{r}, \tilde{p})$. Thus, $|\overline{X}_j| - \tilde{p}_j^- > k - c$. Since

$|\overline{X}_i| - \tilde{p}_i^- \geq |\overline{X}_j| - \tilde{p}_j^-$ for all $i \in N$, then $|\overline{X}_i| - \tilde{p}_i^- > k - c$. Thus,

$$\begin{aligned}
\sum_{i \in N} |\overline{X}_i| - \sum_{i \in N} \tilde{p}_i^- &= \sum_{i \in N} |\overline{X}_i| - \tilde{p}_N \\
&= \sum_{i \in N} |\overline{X}_i| - \tilde{r}_N \\
&= \sum_{i \in N} |\overline{X}_i| - r_N \\
&= nk - \sum_{i \in N} |X_i| - r_N \\
&> nk - nc,
\end{aligned}$$

or equivalently, $(\sum_{i \in N} |X_i| + r_N)/n < c$. Since $c = \min_{i \in N}(|X_i| + p_i^-)$ (by definition) and $c \geq |X_n|$ (by assumption), then it is easy to see that $c = (\sum_{i \in N} |X_i| + r_N)/n$. This is a contradiction since $(\sum_{i \in N} |X_i| + r_N)/n < c$. Thus, $\min_{i \in N} u_i(\tilde{r}, \tilde{p}) \leq \min_{i \in N} u_i(r, p)$. This completes the proof. $\square$

3.4 Simulations

In this section, we evaluate the effect of changing the number of users and number of packets on various statistical parameters in the *random packet distribution* setting. We assume that each packet would be present, independently from other users and other packets, at each user, with probability $0 < p < 1$. All the following simulations assume that the users are following our two algorithms, described previously in this thesis. In the following simulations, we consider three different settings: total of 4 users and 8 packets, total of 6 users and 8 packets, and total of 6 users and 10 packets. We have simulated for the following statistical parameters: $\min_{i \in N}(p_i^-)$, $\max_{i \in N}(p_i^-)$, $\min_{i \in N}(u_i)$, $var_{i \in N}(p_i^+)$, $\sum_{i \in N}(r_i)$, and $avg_{i \in N}(u_i)$.

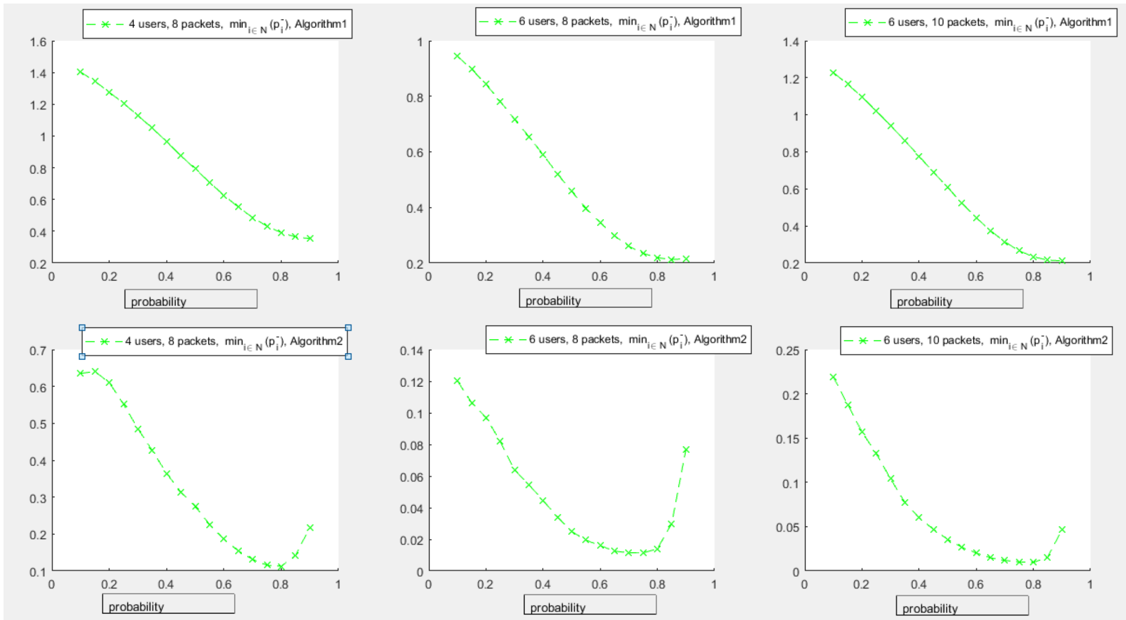

Figure 3.5: $\min_{i\in N}(p_i^-)$

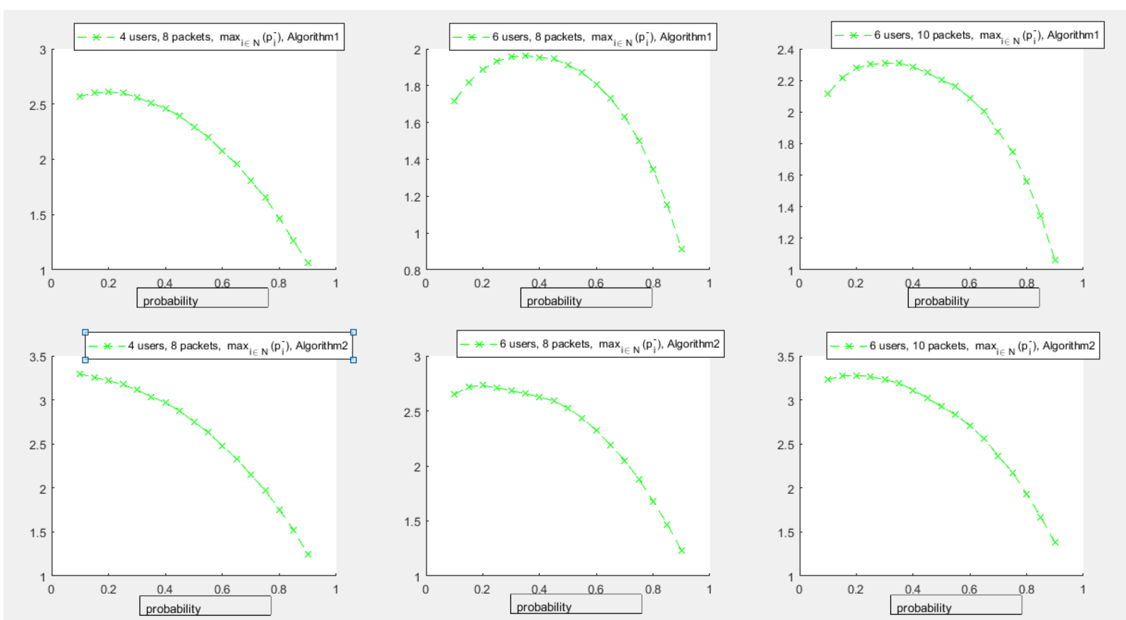

Figure 3.6: $\max_{i\in N}(p_i^-)$

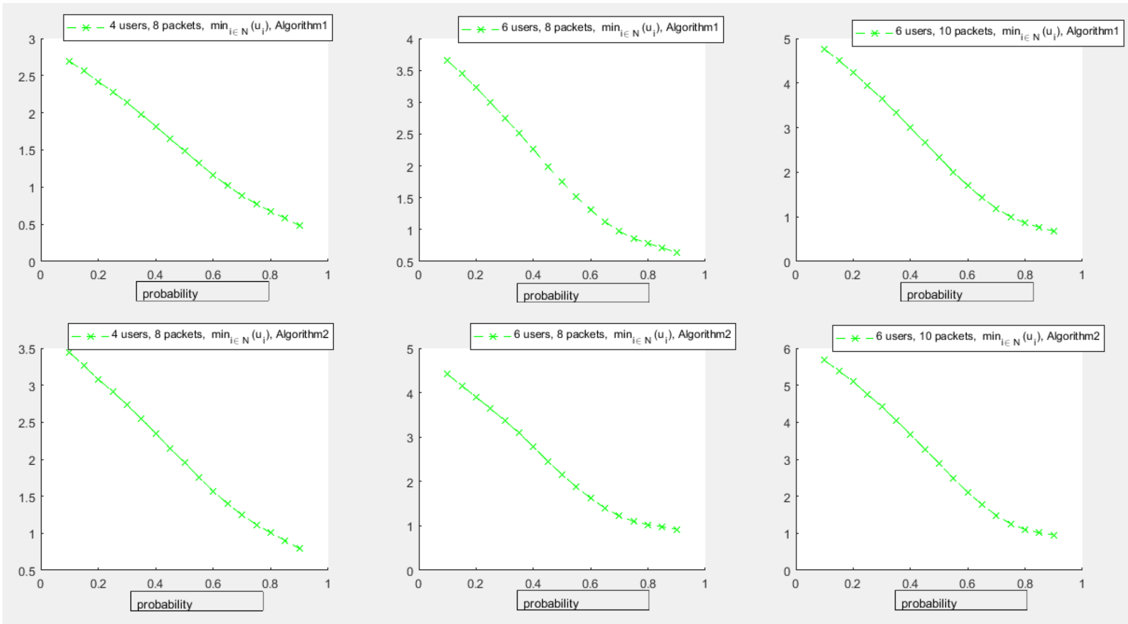

Figure 3.7: $\min_{i\in N}(u_i)$

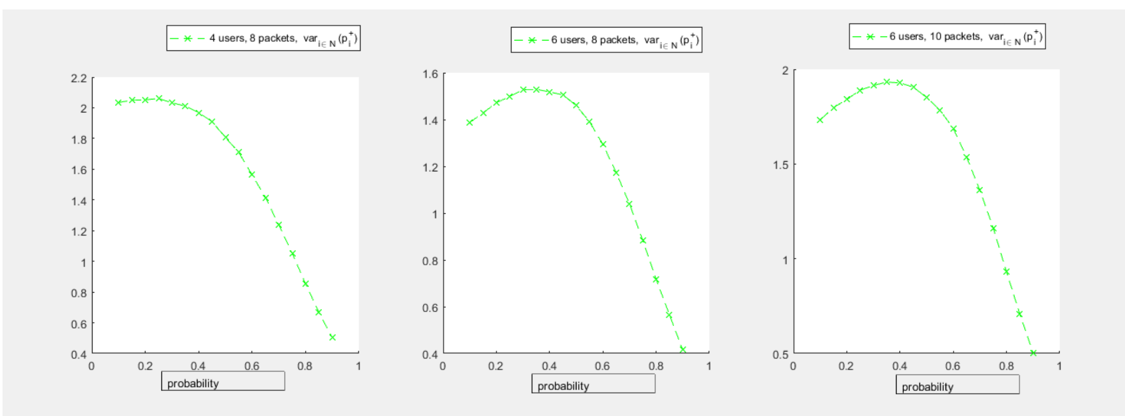

Figure 3.8: $var_{i\in N}(p_i^+)$

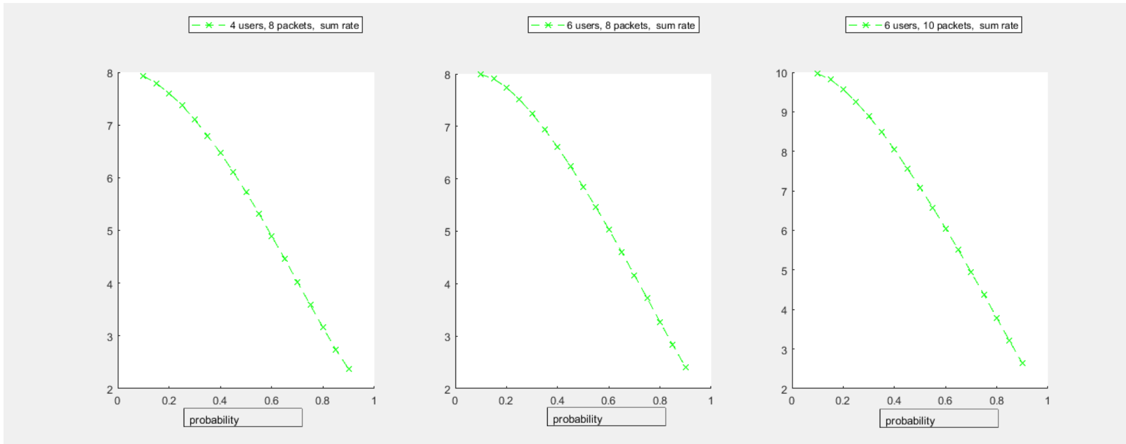

Figure 3.9: $\sum_{i \in N}(r_i)$

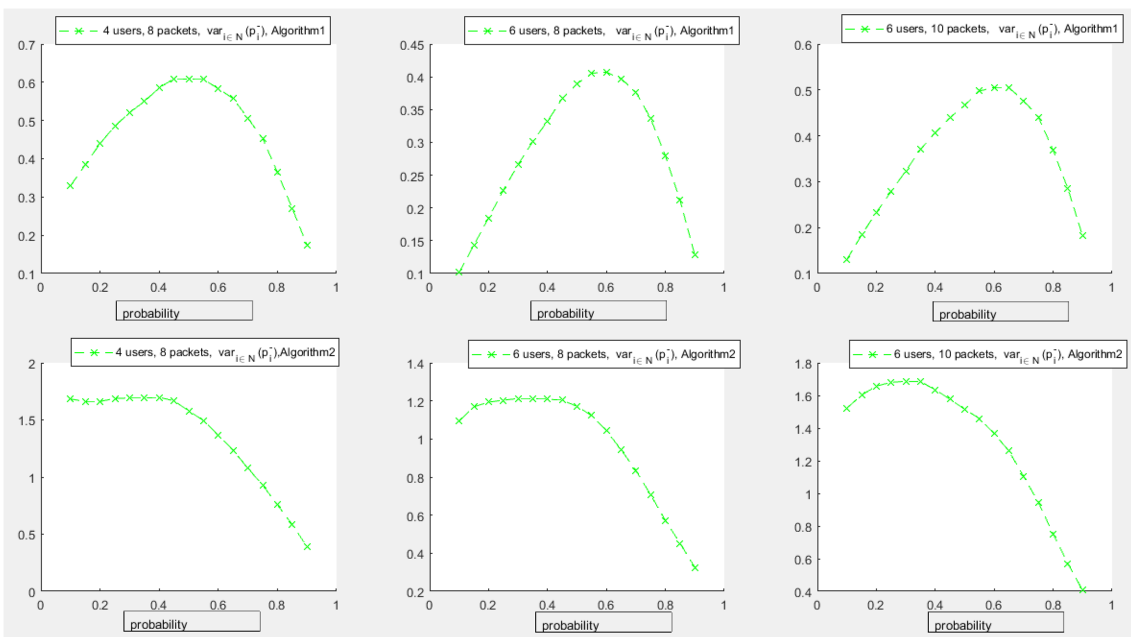

Figure 3.10: $var_{i \in N}(p_i^-)$

We can observe in the graphs corresponding to $\min_{i \in N}(u_i)$ and $\sum_{i \in N}(r_i)$ that the values have a decreasing trend as the value of p increases. This can be attributed to the fact that with more number of packets available with the users from the beginning, there is lesser need of the number of transmissions to make all the users *complete*. Less transmissions eventually result in lesser exchange of money too, which when combined with fewer (r_i), can take a negative hit on the value of (u_i).

We have calculated the 95% confidence intervals for each statistical parameter for each each probability point for the setting where there are a total of 4 users and 8 packets. Below are the values:

For $p = 0.1$,

	Algorithm 1	Algorithm 2
$\max_{i\in N}(p_i^-)$	2.5708 ± 0.00454	3.2924 ± 0.00839
$\min_{i\in N}(p_i^-)$	1.4043 ± 0.00399	0.6383 ± 0.00797
$var_{i\in N}(p_i^-)$	0.33 ± 0.00797	1.6763 ± 0.00395
$\min_{i\in N}(u_i)$	2.6941 ± 0.06766	3.4457 ± 0.00683

$var_{i\in N}(p_i^+)$	2.0279 ± 0.00406
$\sum_{i\in N}(r_i)$	7.9336 ± 0.00406

$\sum_{i\in N}(r_i)$, and $var_{i\in N}(p_i^+)$ are same for both algorithms.

For $p = 0.15$,

	Algorithm 1	Algorithm 2
$\max_{i\in N}(p_i^-)$	2.6034 ± 0.0051	3.2598 ± 0.0089
$\min_{i\in N}(p_i^-)$	1.3419 ± 0.00418	0.6348 ± 0.00791
$var_{i\in N}(p_i^-)$	0.3885 ± 0.00478	1.6712 ± 0.0157

$\min_{i \in N}(u_i)$	2.5544 ± 0.00784	3.2565 ± 0.00714
$var_{i \in N}(p_i^+)$	2.0588 ± 0.0217	
$\sum_{i \in N}(r_i)$	7.7967 ± 0.00607	

For $p = 0.20$,

	Algorithm 1	Algorithm 2
$\max_{i \in N}(p_i^-)$	2.6056 ± 0.00642	3.2119 ± 0.0093
$\min_{i \in N}(p_i^-)$	1.2785 ± 0.00429	0.6126 ± 0.0078
$var_{i \in N}(p_i^-)$	0.4358 ± 0.00564	1.6561 ± 0.0161
$\min_{i \in N}(u_i)$	2.4256 ± 0.00793	3.0842 ± 0.00734
$var_{i \in N}(p_i^+)$	2.0525 ± 0.0217	
$\sum_{i \in N}(r_i)$	7.6064 ± 0.00781	

For $p = 0.25$,

	Algorithm 1	Algorithm 2
$\max_{i \in N}(p_i^-)$	2.5912 ± 0.00729	3.1643 ± 0.00975
$\min_{i \in N}(p_i^-)$	1.2036 ± 0.00439	0.5518 ± 0.00767
$var_{i \in N}(p_i^-)$	0.4824 ± 0.00646	1.6735 ± 0.0163

$\min_{i \in N}(u_i)$	2.2758 ± 0.00802	2.9044 ± 0.00761
$var_{i \in N}(p_i^+)$	2.0719 ± 0.0221	
$\sum_{i \in N}(r_i)$	7.3656 ± 0.00932	

For $p = 0.30$,

	Algorithm 1	Algorithm 2
$\max_{i \in N}(p_i^-)$	2.5604 ± 0.00806	3.1152 ± 0.0102
$\min_{i \in N}(p_i^-)$	1.13 ± 0.00438	0.4868 ± 0.00742
$var_{i \in N}(p_i^-)$	0.5214 ± 0.00736	1.6989 ± 0.0166
$\min_{i \in N}(u_i)$	2.1374 ± 0.00803	2.7339 ± 0.00782
$var_{i \in N}(p_i^+)$	2.0402 ± 0.022	
$\sum_{i \in N}(r_i)$	7.0984 ± 0.0104	

For $p = 0.35$,

	Algorithm 1	Algorithm 2
$\max_{i \in N}(p_i^-)$	2.5208 ± 0.00871	3.0488 ± 0.0107
$\min_{i \in N}(p_i^-)$	$1.0478 \pm 0.0.0136$	0.4188 ± 0.00842

$var_{i\in N}(p_i^-)$	00.5576 ± 0.00803	1.7064 ± 0.0167
$\min_{i\in N}(u_i)$	1.9776 ± 0.0115	2.0087 ± 0.00216
$var_{i\in N}(p_i^+)$	2.0087 ± 0.0216	
$\sum_{i\in N}(r_i)$	6.7999 ± 0.0115	

For $p = 0.40$,

	Algorithm 1	Algorithm 2
$\max_{i\in N}(p_i^-)$	2.4545 ± 0.00935	2.968 ± 0.0111
$\min_{i\in N}(p_i^-)$	$0.9662 \pm 0.0.0436$	0.3693 ± 0.00666
$var_{i\in N}(p_i^-)$	00.5862 ± 0.0088	1.6929 ± 0.0168
$\min_{i\in N}(u_i)$	1.8178 ± 0.00801	2.3507 ± 0.00824
$var_{i\in N}(p_i^+)$	1.9551 ± 0.0255	
$\sum_{i\in N}(r_i)$	6.456 ± 0.0123	

For $p = 0.45$,

	Algorithm 1	Algorithm 2
$\max_{i\in N}(p_i^-)$	2.3939 ± 0.0098	2.8741 ± 0.0113
$\min_{i\in N}(p_i^-)$	$0.8741 \pm 0.0.0428$	0.3202 ± 0.00618

$var_{i\in N}(p_i^-)$	0.6108 ± 0.00932	1.6564 ± 0.0166
$\min_{i\in N}(u_i)$	1.6464 ± 0.00784	1.9037 ± 0.0207
$var_{i\in N}(p_i^+)$	1.9037 ± 0.0207	
$\sum_{i\in N}(r_i)$	6.1070 ± 0.013	

For $p = 0.50$,

	Algorithm 1	Algorithm 2
$\max_{i\in N}(p_i^-)$	2.3028 ± 0.0102	2.7565 ± 0.0115
$\min_{i\in N}(p_i^-)$	$0.7942 \pm 0.0.0409$	0.2733 ± 0.00564
$var_{i\in N}(p_i^-)$	0.6167 ± 0.00988	1.5934 ± 0.0165
$\min_{i\in N}(u_i)$	1.489 ± 0.00755	1.9557 ± 0.00802
$var_{i\in N}(p_i^+)$	1.8152 ± 0.02	
$\sum_{i\in N}(r_i)$	5.7323 ± 0.0134	

Above are the values for 95% confidence interval till $p = 0.5$, similarly we can calculate the values for 95% confidence interval for higher values of p.

4. CONCLUSION AND FUTURE WORK

4.1 Conclusion

We considered the problem of peer-to-peer data exchange in a lossless broadcast setting. Our goal was to ensure the stability of the grand coalition under which it is guaranteed that all users obtain complete information at the end of the exchange process, with there being no incentive for any subset of users to break away on their own. The previous work done with the same motive of stabilizing the CDE had a non-monetary mechanism and was not able to ensure stability in all problem instances. The key novelty of our framework was the design of monetary mechanism that not only ensures stability, but also maximize the social good of the system as a whole. The selfish nature of the users and the exchange of money, in lieu of transmissions, to ensure stability, provides a solid foundation for its pragmatic use in real wireless networks.

4.2 Future Work

Future research work can take multiple directions. In our current CDE setup, we consider only integer packets or a random linear combination of linear packets. It will be interesting to find out a polynomial-time algorithm for CDE. Present literature only provides for a Linear Program to solve a CDE but no algorithm. If successful, we can then try to find a stable solution for grand coalition with fractional packets too. Also, the current problem has a fixed instance and the initial packets corresponding to each user do not change. Further work can be done for a setting where, after a fixed duration of time, a set of packet is added to the present knowledge of each user. This could represent a practical situation where the packets are continuously transferred from the base station to the mobile devices, whereas the mobile devices are also broadcasting coded or uncoded packets among themselves. In the present setup, we have considered the users to be selfish but truthful, where they declare actual set of packets to the other users. This could be changed to a setting where the users are not truthful and an algorithm has to be developed which ensures that being truthful is the best strategy for the users (gives them maximum utility).

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on information theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking (TON)*, vol. 11, no. 5, pp. 782–795, 2003.

[3] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, "The importance of being opportunistic: Practical network coding for wireless environments," 2005.

[4] S. El Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," in *Information Theory (ITW 2010, Cairo), 2010 IEEE Information Theory Workshop on*, pp. 1–5, IEEE, 2010.

[5] S. Biswas and R. Morris, "Opportunistic routing in multi-hop wireless networks," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 1, pp. 69–74, 2004.

[6] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless lans," in *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 121–132, ACM, 2005.

[7] B. N. Karp and H. Kung, *Geographic routing for wireless networks*. Harvard University Cambridge, MA, 2000.

[8] S. Deb, M. Effros, T. Ho, D. R. Karger, R. Koetter, D. S. Lun, M. Médard, and N. Ratnakar, "Network coding for wireless applications: A brief tutorial," IWWAN, 2005.

[9] D. S. Lun, N. Ratnakar, R. Koetter, M. Médard, E. Ahmed, and H. Lee, "Achieving minimum-cost multicast: A decentralized approach based on network coding," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 3, pp. 1607–1617, IEEE, 2005.

[10] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "Xors in the air: Practical wireless network coding," in *ACM SIGCOMM computer communication review*, vol. 36, pp. 243–254, ACM, 2006.

[11] M. Effros, T. Ho, and S. Kim, "A tiling approach to network code design for wireless networks," in *Information Theory Workshop, 2006. ITW'06 Punta del Este. IEEE*, pp. 62–66, IEEE, 2006.

[12] W. Chen, K. B. Letaief, and Z. Cao, "Opportunistic network coding for wireless networks," in *Communications, 2007. ICC'07. IEEE International Conference on*, pp. 4634–4639, IEEE, 2007.

[13] Y.-P. Hsu, N. Abedini, N. Gautam, A. Sprintson, and S. Shakkottai, "Opportunities for network coding: To wait or not to wait," *IEEE/ACM Transactions on Networking (TON)*, vol. 23, no. 6, pp. 1876–1889, 2015.

[14] Y.-P. Hsu and A. Sprintson, "Opportunistic network coding: Competitive analysis," in *Network Coding (NetCod), 2012 International Symposium on*, pp. 191–196, IEEE, 2012.

[15] K.-s. Koo and M. Govindarasu, "Energy-efficient opportunistic network coding algorithms for wireless networks," in *Computer Communication and Networks (ICCCN), 2015 24th International Conference on*, pp. 1–7, IEEE, 2015.

[16] Y. Birk and T. Kol, "Informed-source coding-on-demand (iscod) over broadcast channels," in *Proceedings IEEE INFOCOM*, pp. 1257–1264, IEEE, 1998.

[17] Y. Birk and T. Kol, "Coding on demand by an informed source (iscod) for efficient broadcast of different supplemental data to caching clients," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2825–2830, 2006.

[18] Z. Bar-Yossef, Y. Birk, T. Jayram, and T. Kol, "Index coding with side information," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1479–1494, 2011.

[19] S. Y. El Rouayheb, M. A. R. Chaudhry, and A. Sprintson, “On the minimum number of transmissions in single-hop wireless coding networks,” in *Information Theory Workshop, 2007. ITW’07. IEEE*, pp. 120–125, IEEE, 2007.

[20] M. Langberg and A. Sprintson, “On the hardness of approximating the network coding capacity,” *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1008–1014, 2011.

[21] A. Sprintson, P. Sadeghi, G. Booker, S. Y. El Rouayheb, *et al.*, “A randomized algorithm and performance bounds for coded cooperative data exchange.,” in *ISIT*, pp. 1888–1892, Citeseer, 2010.

[22] A. Sprintson, P. Sadeghi, G. Booker, and S. El Rouayheb, “Deterministic algorithm for coded cooperative data exchange,” in *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, pp. 282–289, Springer, 2010.

[23] N. Milosavljevic, S. Pawar, S. El Rouayheb, M. Gastpar, and K. Ramchandran, “An optimal divide-and-conquer solution to the linear data exchange problem,” in *Proceedings of 2011 IEEE International Symposium on Information Theory (ISIT)*, 2011.

[24] N. Milosavljevic, S. Pawar, S. E. Rouayheb, M. Gastpar, and K. Ramchandran, “Optimal deterministic polynomial-time data exchange for omniscience,” *arXiv preprint arXiv:1108.6046*, 2011.

[25] T. A. Courtade and R. D. Wesel, “Coded cooperative data exchange in multihop networks,” *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1136–1158, 2014.

[26] D. Ozgul and A. Sprintson, “An algorithm for cooperative data exchange with cost criterion,” in *Information Theory and Applications Workshop (ITA), 2011*, pp. 1–4, IEEE, 2011.

[27] S. E. Tajbakhsh, P. Sadeghi, and R. Shams, "A generalized model for cost and fairness analysis in coded cooperative data exchange," in *Network Coding (NetCod), 2011 International Symposium on*, pp. 1–6, IEEE, 2011.

[28] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.

[29] C. Chan *et al.*, *Generating secret in a network*. PhD thesis, Massachusetts Institute of Technology, 2010.

[30] N. Ding, C. Chan, T. Liu, R. A. Kennedy, and P. Sadeghi, "A game-theoretic perspective on communication for omniscience," in *Communications Theory Workshop (AusCTW), 2016 Australian*, pp. 95–100, IEEE, 2016.

[31] A. Heidarzadeh and A. Sprintson, "Cooperative data exchange: A coalition game perspective," in *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*, pp. 1005–1011, IEEE, 2016.