The Index Coding Problem: A Game-Theoretical Perspective

Yu-Pin Hsu, I-Hong Hou, and Alex Sprintson
Department of Electrical and Computer Engineering
Texas A&M University
{yupinhsu, ihou, spalex}@tamu.edu

Abstract—The Index Coding problem has recently attracted a significant interest from the research community. In this problem, a server needs to deliver a set of packets to a group of wireless clients over a noiseless broadcast channel. Each client requests a subset of packets and has another subset given to it as side information. The objective is to satisfy the demands of all clients with the minimum number of transmissions.

In this paper, we study the Index Coding problem from the game-theoretic perspective. We assume that each client is selfish and has a hidden private value for each packet it requests. The objective of the server is to maximize the value of social welfare that captures the trade-off between values of the transmitted packets and the transmission cost incurred by the server. The transmission process is decided through an auction in which the clients are required to submit bids to the server.

Our goal is to design a truthful auction scheme that provides an incentive for each client to bid the true value of the packets and maximizes the value of the social welfare. The key challenge in this context is to determine the encoding functions of the transmitted packets. Since finding an optimal encoding function is an \( N P \)-hard problem, we propose efficient algorithms that identify the encoding functions as well as a payment scheme that provide an approximate solution and guarantee truthfulness.

I. INTRODUCTION

The Index Coding problem is one of the basic problems in wireless network coding. Recently, this problem has attracted a significant interest from the research community (see e.g., [1, 2]). An instance of the Index Coding problem includes a server, a set of wireless clients, and a set \( P = \{p_1, \ldots, p_m\} \) of \( m \) packets that need to be delivered to clients. Each client is interested in a certain subset of packets in \( P \) and has a (different) subset of packets given to it as side information. The server can transmit the original packets or their combinations to clients via a noiseless broadcast channel. The goal is to find a transmission scheme that requires the minimum number of transmissions to satisfy the requests of all clients.

Fig. 1 depicts an instance of the Index Coding problem, where a server needs to deliver four packets \( P = \{p_1, \ldots, p_4\} \) to four clients. It can be verified that the demands of all clients can be satisfied by broadcasting three packets: \( p_1 + p_2, p_2 + p_3, \) and \( p_4 \) (all operations are over \( GF(2) \)). Note that the traditional approach (without coding) requires the transmissions of all four packets \( p_1, \ldots, p_4 \).

The prior works on the Index Coding problem have focused on developing algorithms, establishing rate bounds, and analyzing the computational complexity of the problem [3–8]. In contrast to the original Index Coding problem where the goal of the server is to satisfy the demands of all clients, we focus on settings where the server’s decisions are driven by the values of the packets transmitted to the clients and by the transmission cost. This is a new problem in the intersection of the coding theory and game theory and is related to the framework of mechanism design with social choices [9].

Intuitively, since each transmission at the server incurs a certain cost, the server may decide to transmit a packet only when the packet is important enough for the clients. Thus, our goal is to maximize the social welfare, i.e., the total value of the packets the clients are able to decode minus the total cost of transmitting the packets by the server. The social welfare reflects the positive externalities in terms of clients’ valuations of packets and the negative externalities related to the transmission costs of server.

To calculate the social welfare, the server needs to know, for each client in the system, the client’s valuation of all packets it requests. This information is usually private and selfish clients could not be expected to reveal this information to the server. Accordingly, we aim to design a tractable auction mechanism that provides an incentive for each client to reveal the true
valuation of its packets.

As part of the auction, the clients submit bids to the server, each bid specifies the amount of money the client is willing to pay in order to obtain a packet it wants. Based on the auction mechanism, the server decides whether to accept or reject the bids. The auction mechanism includes two main components, namely the coding algorithm and payment function. The coding algorithm determines which packets are transmitted over the channel and how they are encoded. The payment function determines the amount of money the client will pay to the server for each packet it is able to decode.

Contributions. A classical solution for the problem of maximizing social welfare in the presence of selfish clients is to use the Vickery-Clarke-Groves (VCG) mechanism [10–12]. Such mechanism was shown to be truthful in the sense that each client maximizes its own utility by bidding its true valuations of packets. Unfortunately, finding VCG based encoding functions for our problem is an intractable ($NP$-hard) optimization problem. Accordingly, we present an alternative auction mechanism which is both truthful and computationally efficient. Our mechanism includes a non-VCG based payment scheme that achieves the truthfulness property in the presence of the approximation solution for the problem.

II. Model

We begin the problem with the definition of the Index Coding problem. The input to the problem includes a server $S$, a set of $n$ wireless clients $\Lambda = \{c_1, \ldots, c_n\}$, and a noiseless wireless broadcast channel. The server has a set of $m$ packets $P = \{p_1, \ldots, p_m\}$ that need to be distributed to clients in $\Lambda$. We assume, without loss of generality, that each client requests a single packet in $P$ (i.e., $n = m$) and has access to a subset of packets in $P$ as side information. Indeed, a client that wants more than one packet can be substituted by multiple clients that share the same side information set. In particular, each client $c_i \in \Lambda$ is characterized by the ordered pair $(w_i, H_i)$, where $w_i \in P$ is the packet requested by $c_i$ and $H_i \subseteq P$ is the set of packets available to $c_i$ as side information. We assume that each packet $p_i$ represents an element of the Galois field $GF(q)$.

We assume that each client $c_i \in \Lambda$ has the internal (private) valuation $v_i$ for the packet $w_i$ it requests. The transmission process includes an auction, in which each client submits a bid $b_i$ for packet $w_i$. We denote by $V = \{v_1, \ldots, v_n\}$ and $B = \{b_1, \ldots, b_n\}$ the arrays that include the internal valuations and bids of the clients. Based on the bid vectors, the server identifies linear combinations that will be transmitted over the channel. In this paper, we focus on the scalar-linear coding schemes in which each packet $q_i$ transmitted by the server at the iteration $i$, $1 \leq i \leq \eta$, is a linear combination of packets in $P$. Here, $\eta$ is the total number of transmissions made by the server and $g_i = [g_{i,1}, \ldots, g_{i,m}] \in GF(q)^m$ is the encoding vector for the iteration $i$. We denote by $\Gamma = [g_i]$ the encoding matrix whose rows correspond to encoding vectors of the packets transmitted over the channel. A sparse code [13] is the linear code in which each transmitted packet from the server is a linear combination of at most two packets in $P$. Without loss of generality, we assume that the cost of transmitting a packet by the server is equal to one, hence the total cost incurred by the server is equal to $\eta$.

The goal of the server is to maximize the social welfare, defined as

$$\sum_{i=1}^{n} v_i \cdot \theta_i - \eta,$$

where $\theta_i$ is the indicator function that specifies whether client $c_i$ is able to decode the required packet $w_i$. In particular, $\theta_i = 1$ if there exists a linear decoding function $r_i$ such that $w_i = r_i(g_{i,1}, \ldots, g_{i,m}, H_i)$; otherwise, $\theta_i = 0$. We denote by $\Theta = \{\theta_1, \ldots, \theta_n\}$ the vector of indicator functions.

For example, in the setting depicted in Fig. 1, the social welfare of broadcasting the solution to the Index Coding problem, i.e., $\{p_1 + p_2; p_2 + p_3; p_4\}$, is $0.2 + 0.9 + 0.5 + 0.6 - 3 = -0.8$, while the social welfare of the sequence consisting of a single combination $\{p_3 + p_4\}$ has a higher value of $0.5 + 0.6 - 1 = 0.1$. Thus, transmitting a single combination $p_3 + p_4$ would be more desirable than transmitting three packets $\{p_1 + p_2; p_2 + p_3; p_4\}$ from the server’s point of view.

The payment function is an important part of the auction mechanism. Server determines the amount of payment $\phi_i$ that each client needs to pay for the obtained packet. We denote by $\Phi = \{\phi_1, \ldots, \phi_n\}$ the payment vector for all clients. The value of $\phi_i$ is a function of $B$ and the encoding matrix $\Gamma$ (the latter determines the vector of indicator functions $\Theta$). We assume that the auction mechanism, including the algorithm for determining the encoding function and the payment function is known to all the parties.

We also assume that every client $c_i$ is selfish and chooses its best bidding policy $b_i$ that maximizes its utility defined by $u_i(b_i, B_{-i}) = v_i \cdot \theta_i - \phi_i$, where $B_{-i} = B \setminus \{b_i\}$. Note that $u_i(b_i, B_{-i})$ depends on the bids of all clients as well as the server’s mechanism of choosing the encoding matrix and the payment function. We say that a mechanism is truthful if $u_i(v_i, B_{-i}) \geq u_i(b_i, B_{-i})$ for all $b_i$ and $B_{-i}$. That is, regardless of the bids submitted by other clients, the utility of client $c_i$ is maximized when $b_i = v_i$.

In this paper, we design the encoding matrix and the payment function that satisfies the following two conditions: (i) every client is incentivized to report its true valuation of the desired packet (i.e., truthful property) and (ii) the social welfare is maximized.

III. VCG-BASED MECHANISM DESIGN

In this section, we present a truthful mechanism based on the celebrated Vickrey-Clarke-Groves (VCG) approach [10–12].

Our scheme uses the following objective function

$$w(B, \Gamma) = \sum_{i=1}^{n} b_i \cdot \theta_i - \eta.$$
Note that this function is similar to the social welfare, but instead of vector of bids $B$ we use a vector of packet valuations $V$. In addition, we define $w_{-i}(B, \Gamma) = \sum_{j \neq i} b_j \cdot \theta_j - \eta$. In our mechanism, the server chooses the encoding matrix that maximizes the value of $w(B, \Gamma)$ for the given bids $B$.

**VCG-coding scheme:** Choose the encoding matrix $\Gamma^*$ that maximizes the value of $w(B, \Gamma)$, i.e.,

$$\Gamma^* = \arg \max_{\Gamma} w(B, \Gamma). \quad (1)$$

If there are multiple encoding functions that maximize $w(B, \Gamma)$, we choose one that satisfies the larger number of clients. Note that without loss of generality, we can assume that $v_i \in [0, 1]$, since client $c_i$ is assured to get the desired packet when submitting the bid $b_i = 1$.

Our scheme uses the following payment function $\phi_i$:

**VCG-payment function:** The client $c_i$ is charged as follows.

$$\phi_i = \max_{\Gamma} w_{-i}(B, \Gamma) - w_{-i}(B, \Gamma^*). \quad (2)$$

The first term in Equation (2) corresponds to the maximum value of the objective function when client $c_i$ is removed, while the second term represents the objective function for all clients excluding $c_i$ when the optimal encoding matrix $\Gamma^*$ determined by Equation (1) is employed. The VCG-payment function charges the “externality” of client $c_i$, i.e., the decrease in optimal objective function when client $c_i$ is included in the system.

**Proposition 1.**

1) The VCG-coding scheme associated with the VCG-payment function is a truthful mechanism, i.e., $B = V$. Moreover, the pricing function in Equation (2) is non-negative, and the utilities of all clients are non-negative, i.e., $\phi_i \leq v_i$.

2) The VCG-coding scheme associated with the VCG-payment function maximize the value of social welfare.

The proof of Proposition 1 follows from the properties of the VCG mechanism [9] and is omitted due to space constraints.

The problem of finding an optimal VCG encoding matrix $\Gamma^*$ is an interesting combinatorial problem by itself. Unfortunately, this problem is computationally intractable. More specifically, for the instance of the problem in which the bid of each client $b_i$ is equal to one, there exists an optimal solution $\Gamma^*$ that satisfies all clients, since the client’s bid is equal to the cost of transmitting a single packet. Then, the problem of maximizing the value of $w(B, \Gamma)$ (where $w(B, \Gamma) = n - \eta$) is equivalent to the problem of minimizing the required number of transmissions to satisfy all clients, which is $\mathcal{NP}$-hard.

**IV. APPROXIMATION ALGORITHM**

In this section, we present an approximation algorithm for the special multiple unicast case, in which every packet is required by exactly one client. We also focus on finding a solution that uses sparse codes. With sparse coding, both encoders and decoders can be implemented in a simpler manner which has a significant advantage for practical applications.

Our scheme uses the notion of a weighted dependency graph described below.

**Definition 2.** The weighted dependency graph is a directed graph $G(V, A)$ defined as follows:

- For each client $c_i \in \Lambda$, there is a corresponding vertex $\zeta_i$ in $V$.
- For any two clients $c_i$ and $c_j$ such that $w_i \in H_j$, there is a directed arc $(\zeta_i, \zeta_j) \in A$.
- The weight $\gamma_a$ of arc $a = (\zeta_i, \zeta_j)$ is equal to $b_i$.

Fig. 2 illustrates the weighted dependency graph $G(V, A)$ corresponding to the instance of the problem depicted in Fig. 1. Note that for each cycle $C \in G(V, A)$, the server can save one transmission. That is, for a cycle $C$ of length $|C|$, all of the clients that correspond to its vertices can be satisfied using $|C| - 1$ transmissions. In particular, the clients that correspond to the vertices of $C$ share the transmission cost of $|C| - 1$. In the example depicted in Fig. 2 we have a cycle $(\zeta_1, \zeta_2, \zeta_3)$ that involves three clients. It is easy to verify that these three clients can be satisfied by two transmissions: $p_1 + p_2$ and $p_2 + p_3$. Note that if there is no cycle in $G(V, A)$ then a sparse code cannot improve the value of $w(B, \Gamma)$.

Note that reduced transmission cost results in a higher value of the social welfare. Unfortunately, even with sparse coding, the problem of finding an optimum encoding function under the VCG scheme is intractable.

**Lemma 3. In multiple unicast scenario the problem of finding the a sparse code $\Gamma$ that maximizes the value of $w(B, \Gamma)$ is $\mathcal{NP}$-hard.**

The proof of Lemma 3 uses the reduction from the cycle packing problem and follows the same lines as the complexity proof in [13]. The details are omitted due to the space constraints.

**A. Approximation Algorithm**

In this section, we present an alternative mechanism that uses Algorithm 1 to determine encoding functions and Algorithm 2 to determine clients payments. The main idea of our approach is to find an approximation solution to the problem of maximizing $w(B, \Gamma)$ and then use a pricing scheme that uses this approximation solution to motivate all clients to reveal the true valuations of the required packets.

We define the weight of cycle $C$ in $G(V, A)$ as

$$\gamma(C) = \sum_{a \in C} \gamma_a - (|C| - 1),$$

Fig. 2. Weighted dependency graph for the instance of the multiple unicast problem in Fig. 1.
Algorithm 1:

\[ \text{input : Bids vector } \mathbf{b} \text{ and side information } H, \text{ for all clients} \]
\[ \text{output: Encoding matrix } \Gamma^* \]
1. Construct the weighted dependency graph \( G(V, A) \);
2. Define the cost \( \lambda_a \) of arc \( a \in A \) by \( \lambda_a = 1 - \gamma_a \);
3. \( S_1, S_2 \leftarrow \emptyset \)
4. while There is a cycle in \( G(V, A) \), with the cost less than or equal to one do
5. \( \text{Find the cycle } C \text{ in } G(V, A) \text{ with the smallest cost}; \)
6. \( \text{Find } |C| - 1 \text{ linear combinations that satisfy clients corresponding to vertices in } C; \)
7. Add \( |C| - 1 \) corresponding vectors as new rows in \( \Gamma^* \); \( S_1 \leftarrow S_1 \cup \{ c_i : \zeta_i \in C \} \);
8. Remove all vertices of \( C \), along with incident arcs, from the graph.
9. end
10. for \( i \leftarrow 1 \text{ to } n \) do
11. \( \text{if } b_i = 1 \text{ but } c_i \notin S_1 \text{ then} \)
12. \( \text{Add a singleton vector that corresponds to } w_i \text{ as a new row in } \Gamma^*; \)
13. \( S_2 \leftarrow S_2 \cup \{ c_i \}; \)
14. end
15. end

where the first term is the sum of all bids and the second term is the total transmission cost of all packets that correspond to the arcs of this cycle.

Let \( C \) be a set of vertex-disjoint cycles in \( G(V, A) \). Then, let \( \Gamma \) be the encoding function that includes \( \sum_{C \in C} (|C| - 1) \) linear combinations that correspond to the cycles in \( C \) (\( |C| - 1 \) linear combinations for each cycle \( C \in C \)). Note that \( \Gamma \) satisfies all clients that correspond to the vertices of \( C \). Note also that
\[ w(B, \Gamma) = \sum_{C \in C} \gamma(C). \]

The idea of Algorithm 1 is to iteratively identify and remove a maximum weight cycle. We note that \( \gamma(C) \) can be written as
\[ \gamma(C) = 1 - \sum_{a \in C} (1 - \gamma_a). \]
Thus, in this case, a maximum weight cycle with respect to weight \( \gamma(C) \) corresponds to the minimum cost cycle with respect to cost \( \sum_{a \in C} (1 - \gamma_a) \). Therefore, we define the cost of arc \( a \in A \) by \( \lambda_a = 1 - \gamma_a \) and the cost of cycle \( C \) by
\[ \lambda(C) = \sum_{a \in C} \lambda_a. \]
The algorithm finds the minimum cost cycle in line 5. To this end, we can use well-known polynomial time algorithms such as Floyd-Warshall algorithm. After such a cycle is identified, it is removed from the graph in line 9. The condition in line 4 guarantees that the maximum weight cycle in the remaining graph \( G(V, A) \) (i.e., the minimum cost cycle) has a non-negative weight. Line 13 ensures that the clients whose bid is equal to the transmission cost are served. Such clients are then included in set \( S_2 \). Set \( S_2 \) contains the clients that correspond to the cycles chosen in line 5. Both set \( S_1 \) and \( S_2 \) will be used by Algorithm 2 in order to determine the payments of the clients. Note that all procedures in Algorithm 1 require a polynomial number of steps, hence Algorithm 1 can be executed in polynomial time.

Unfortunately, the approximate solution \( \Gamma^* \) identified by Algorithm 1 cannot be used for the VCG-payment functions specified by Equation (2). In particular, the following example shows that such combination might not satisfy the truthfulness property.

**Example 4.** Consider the instance of the problem depicted in Fig. 3. We show that client \( c_1 \) does not have an incentive to reveal the true valuation of its packets. Suppose that client \( c_1 \) submits the bid \( b_1 = 0.55 \) equal to its internal value of packet \( p_1 \). In this case, Algorithm 1 will identify the cycle \((c_2, c_3)\) and output the corresponding solution \( p_2 + p_3 \). In this case, the utility \( u_1 \) of \( c_1 \) is equal to zero.

Now, suppose that \( c_1 \) submits bid \( b_1 = 0.7 \). In this case, Algorithm 1 will transmit the sequence of packets \( \{p_1 + p_2 : p_3 + p_4\} \) that corresponds to cycles \((c_1, c_2)\) and \((c_3, c_4)\), respectively. According to Equation (2), the payment of client \( c_1 \) is equal to \( \phi_1 = (0.6 + 0.6 - 1) - (0.6 + 0.6 + 0.55 - 2) = 0.45 \). In this case, the utility of client \( c_1 \) is equal to 0.1, which is bigger than the case in which the client is truthful.

**B. Payment scheme**

In this section, we propose a payment scheme for our algorithm. We start with a definition. We say that a coding algorithm is *monotone* if, for any \( B^- \), there exists a single critical value \( \theta_1 \) such that \( c_i \) gets the desired packet (i.e., \( \theta_1 = 1 \)) when \( b_i \geq \theta_1 \); otherwise, \( c_i \) does not get the packet it needs.

Using the results of Mualem and Nisan [14] it can be shown that if the coding mechanism is monotone and the payment scheme is based on the critical value (i.e., each client is required to pay the critical value \( \theta_1 \)), then the resulting mechanism is truthful. We prove in Lemma 5 below that Algorithm 1 is monotone. Accordingly, we use Algorithm 2 (presented below) to compute the critical value for each client.
Algorithm 2:

input: Bids vector B; side information H_i for all clients; sets S_1 and S_2 returned by Algorithm 1; client c_i
output: Payment φ_i for client c_i
1. If c_i ∉ S_1 ∪ S_2 return φ_i = 0;
2. If c_i ∈ S_2 return φ_i = 1;
3. Construct the weighted dependency graph G(V, A) with the following costs:
   - for arc (i, j) ∈ A, λ_{ij} = 1
   - for arc (i, j) ∈ A, λ_{ij} = 1 - γ_{ij}
4. β ← ∞
5. while there exists a cycle C in G(V, A) such that λ(C) is less than or equal to one, as well as there exists a cycle that includes vertex \( \zeta_i \) do
   - Find the cycle C��rst in G(V, A) with the smallest cost.
   - Find the cycle C_2 in G(V, A) that traverses vertex \( \zeta_i \) and has the smallest cost among those that go through vertex \( \zeta_i \); if \( λ(C_2) - λ(C_1) = 0 \) then
     - return φ_i = 0
   - end else
     - β ← min{β, λ(C_2) - λ(C_1)}
     - Remove all vertices of C_1 along with the arcs incident to them from G(V, A)
   - end
6. Find a cycle C_2 in G(V, A) (if any) that traverses vertex \( \zeta_i \) and has the smallest cost among these that go through vertex \( \zeta_i \);
7. β ← min{β, λ(C_2) - 1}
8. return φ_i = β;

Algorithm 2 is similar to Algorithm 1, but has a different cost function. To compute the payment for client c_i ∈ S_1, we assign the unit cost for every outgoing arc \((\zeta_i, \zeta_j)\) of vertex \( \zeta_i \) (i.e., set \( γ_{ij} = 0 \) and \( b_i = 0 \)). For each iteration of Line 5, we calculate the difference of the cost of cycles C_1 and C_2 in G(V, A), where C_1 has the minimum global cost while C_2 is the local optimal cycle that traverses vertex \( \zeta_i \). In other words, given B_{−i}, we are looking for the lowest bid \( b_i \) such that a cycle containing vertex \( \zeta_i \) in the remaining G(V, A) would be selected by Line 5 of Algorithm 1.

The value β is equal to the minimum value of \( λ(C_2) - λ(C_1) \) among all iterations. The payment of each client in S_1 will be equal to β.

C. Analysis

In this section we prove that Algorithms 1 and 2 result in a truthful mechanism. We begin by showing that Algorithm 1 is monotone.

Lemma 5. Algorithm 1 is monotone.

Proof: (Sketch) Suppose client c_i is able to recover the desired packet. Then, a cycle containing vertex \( \zeta_i \) is chosen by Algorithm 1. If c_i increases its bid, all cycles that contain \( \zeta_i \) will have lower cost. Therefore, the algorithm will select a cycle containing vertex \( \zeta_i \) and the lemma follows.

Lemma 6. Given B_{−i}, for each node c_i that is able to decode the required packet, the payment function φ_i is equal to the critical value.

Proof: (Sketch) At each iteration of Algorithm 2, we calculate the lowest bid \( b_i \) for client c_i such that a cycle containing vertex \( \zeta_i \) would be selected by Algorithm 1, i.e., client c_i can get the desired packet w_i.

We summarize our results in the following theorem.

Theorem 7. The Algorithms 1 and 2 run in polynomial time and result in a truthful mechanism.

Proof: It is easy to verify that the both algorithms require a polynomial number of steps. The truthfulness property follows from lemmas 5, 6, and [14, Theorem 1].

V. Conclusion

In this paper, we focus on a new challenging problem that lies in the intersection of game theory and coding theory. Based on the Vickrey-Clarke-Groves (VCG) mechanism, we propose the VCG-coding scheme and the VCG-payment function that maximize the social welfare and provide incentives for the clients to submit truthful bids. Unfortunately, in many cases of practical interest it is intractable to implement the exact VCG-coding scheme. Accordingly, we proposed an approximation algorithm as well as the corresponding payment scheme that guarantee trustfulness.

References