Algorithmic View of Online Prize-collecting Optimization Problems

Christine Markarian[1] and Abdul Nasser El-Kassar[2]

[1]*Department of Engineering and Information Technology, University of Dubai, U.A.E.*

[2]*Department of Information Technology and Operations Management, Lebanese American University, Lebanon*

Keywords: Prize-collecting, Facility Location Planning, Decision Making, Online Algorithms, Competitive Analysis.

Abstract: *Online algorithms* have been a cornerstone of research in network design problems. Unlike in classical *offline algorithms*, the input to an online algorithm is revealed in portions over time and the online algorithm reacts to each portion while targeting a given optimization goal. Online algorithms are deployed in real-world optimization problems in which provably good decisions are expected in the present without knowing the future. In this paper, we consider a well-established branch of online optimization problems, known as *online prize-collecting* problems, in which the online algorithm may reject some input portions at the cost of paying an associated penalty. These appear in business applications in which a company decides to lose some customers by paying an associated penalty. Particularly, we study the online prize-collecting variants of three well-known optimization problems: Connected Dominating Set, Vertex Cover, and Non-metric Facility Location, namely, *Online Prize-collecting Connected Dominating Set* (OPC-CDS), *Online Prize-collecting Vertex Cover* (OPC-VC), and *Online Prize-collecting Non-metric Facility Location* (OPC-NFL), respectively. We propose the first online algorithms for these variants and evaluate them using *competitive analysis*, the standard framework to measure online algorithms, in which the online algorithm is measured against the optimal offline algorithm that knows the entire input sequence in advance and is optimal.

1 INTRODUCTION

Many real-world optimization problems are *online* in nature, requiring smart decisions to be made immediately, even when future information is not fully known. The challenge is then to make decisions that will cause as few regrets as possible in the future. Classically, most optimization problems have been modeled in an *offline* setting, in which the entire input sequence is assumed to be known to the decision maker in advance. In many real-world scenarios, having access to future input sequences is most of the time hard or even impossible. Over the past decades, *online algorithms* have gained a lot of popularity in such scenarios, since they can provide performance guarantee to the decision maker. That is, decisions made by online algorithms are measured against optimal decisions that are made by the assumption of knowing the entire input sequence in advance. Hence, a decision maker would know how good or bad the decision is at the time it is made. Moreover, it would allow decision makers to know and try to achieve the best provably possible decision achievable by any decision maker. Unlike in classical offline models, the input to an online algorithm is revealed in portions over time and the online algorithm reacts to each portion, as it targets a given optimization goal against the entire input sequence.

We adopt the *competitive analysis* framework to evaluate the online algorithms (Sleator and Tarjan, 1985). In this framework, the performance of the online algorithm is measured against the optimal offline algorithm, that knows the entire input sequence in advance and is optimal, in the worst case. Given an input sequence σ, let $\mathcal{C}_A(\sigma)$ and $\mathcal{C}_{OPT}(\sigma)$ denote the cost incurred by an online algorithm A, possibly randomized, and an optimal offline algorithm OPT, respectively. A has *competitive ratio* c or is *c-competitive* if there exists a constant α such that $\mathcal{C}_A(\sigma) \leq c \cdot \mathcal{C}_{OPT}(\sigma) + \alpha$ for all input sequences σ. We assume the *oblivious adversarial model*, in which the adversary specifies all of the input at the beginning and does not know the random outcomes of the algorithm.

In this paper, we consider a well-established branch of online optimization problems known as *online prize-collecting* problems, in which the online algorithm may reject some input portions at the cost of paying an associated penalty. These appear in network planning applications for service providers, in

Markarian, C. and El-Kassar, A.
Algorithmic View of Online Prize-collecting Optimization Problems.
DOI: 10.5220/0010471507440751
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 1*, pages 744-751
ISBN: 978-989-758-509-8

which a company may decide to lose some customers by paying a corresponding penalty in the revenue. Since (Qian and Williamson, 2011) introduced the online prize collecting framework, many graph optimization problems, such as variants of Steiner problems and metric Facility Location, were defined in this model (see (Felice et al., 2015; Hajiaghayi et al., 2013; Markarian, 2018)). It is worth noting that the online prize-collecting model is a generalization of the online model in which all penalties are set to infinity.

2 OUR CONTRIBUTION

In this paper, we study the online prize-collecting variants of three classical optimization problems: Connected Dominating Set, Vertex Cover, and Nonmetric Facility Location.

2.1 Online Prize-collecting Connected Dominating Set

Consider an advertising company trying to reach potential customers in a social network for the purpose of advertising a certain product or service. The company assumes a customer is reached either if he has direct access to the ad or one of his friends has. Every now and then, a number of potential customers are announced. These are people who would likely be interested in the ad. The point is to send the ad to as few people as possible so as to reduce costs, while reaching as many people as possible. To achieve this, the company wishes to find a group of connected people who can spread the message to each other and to the rest of the potential customers through the word-of-mouth effect. The company may ignore some of the potential customers, by paying an associated penalty. The goal is to minimize the total costs of sending ads *and* penalties.

From an algorithmic perspective, this scenario can be formalized as the *Online Prize-collecting Connected Dominating Set* problem (OPC-CDS). OPC-CDS is the online prize-collecting variant of Connected Dominating Set. It generalizes the *Online Set Cover* problem (OSC) introduced by (Alon et al., 2003), defined as follows.

Definition 1. *(OSC) Given a universe $\mathcal{U}$ of elements and a collection $\mathcal{S}$ of subsets of $\mathcal{U}$, each associated with a cost. A subset $D \subseteq \mathcal{U}$ of elements arrives over time and OSC asks to find a minimum cost of subsets $\mathcal{C} \subseteq \mathcal{S}$ that cover all elements in D.*

(Korman, 2005) gave a lower bound of $\Omega(\log m \log n)$ on the competitive ratio of any online polynomial-time randomized algorithm for OSC, under the assumption that NP $\nsubseteq$ BPP, where m is the number of subsets and n is the number of elements. This implies a lower bound of $\Omega(\log^2 n)$ on the competitive ratio of any randomized polynomial-time algorithm for OPC-CDS, where n is the number of nodes. OPC-CDS is defined as follows.

Definition 2. *(OPC-CDS) Given an undirected connected graph $G = (V,E)$ with $|V| = n$, node-weight function $w : V \to \mathcal{R}^+$, and penalty-cost function $p : V \to R^+$. A sequence of disjoint subsets of V arrives over time. A subset $S \subseteq V$ serves as a connected dominating set of a given subset $D \subseteq V$ if every node in D is either in S or has an adjacent node in S, and the subgraph induced by S is connected in G. In each step t, a subset $D_t \subseteq V$ arrives; for each $u \in D_t$, OPC-CDS asks to either pay the penalty p_u of u or add u to a subset $D'_t \subseteq D_t$ that is served by a connected dominating set, at time t. The goal is to minimize the total weight of the connected dominating set constructed and the total penalties paid.*

To the best of our knowledge, no online algorithm with non-trivial competitive ratio exists for OPC-CDS. A special case of OPC-CDS, in the online model, namely, the *Online Connected Dominating Set* problem (OCDS), was introduced by (Hamann et al., 2018), in the context of modern robotic warehouses. Unlike in this paper, (Hamann et al., 2018) studied a special case in which all node-weights are uniform. Hence, their approach cannot be applied to our problem.

In this paper, we propose the first online algorithm for OPC-CDS, with $O(\frac{w_{max}}{w_{min}} \log^2 n)$-competitive ratio, where n is the number of nodes, w_{max} is the maximum node weight, and w_{min} is the minimum node weight. Our algorithm is randomized and makes use of the deterministic algorithm of (Alon et al., 2003) for the *Online Set Cover* problem (OSC), defined earlier, and the randomized algorithm of (Hajiaghayi et al., 2014) for the *Online Node-weighted Steiner Tree* problem (OPC-NWST).

2.2 Online Prize-collecting Vertex Cover

Consider the advertising company described earlier. For some ads, the company wishes to assure that the ad is reached to the customers without relying on the word-of-mouth effect. This means, it assumes a customer is influenced only through close friends. Every now and then, a number of friendships are announced.

These are pairs of people who would likely influence each other. If one of them receives the ad, the other is assumed to have been reached. Moreover, the company may wish, as before, to ignore some friendships, at the cost of paying a penalty. The goal is to minimize the total costs of sending ads *and* penalties.

This scenario can be modeled as the *Online Prize-collecting Vertex Cover* problem (OPC-VC). OPC-VC is the online prize-collecting variant of Vertex Cover and is defined as follows.

Definition 3. *(OPC-VC) Given an undirected graph $G = (V,E)$ with $|V| = n$ and node-weight function $w : V \to \mathcal{R}^+$. A sequence of edges, each associated with a penalty, arrives over time. In each step, an edge arrives; OPC-VC asks to output a set S of nodes, such that each edge has at least one of its endpoints in S or its penalty is paid, at the current step. The goal is to minimize the total weight of S and the total penalties paid.*

To the best of our knowledge, no online algorithm with non-trivial competitive ratio exists for OPC-VC. A special case of OPC-VC is the *Online Vertex Cover* problem (OVC). For the unweighted variant of OVC, in which all node weights are uniform, there is a simple greedy algorithm with 2-competitive ratio. As soon as an edge arrives, if it is not covered, this algorithm adds both of its endpoints to the solution.

In this paper, we propose the first online algorithm for OPC-VC, with 3-competitive ratio. Our algorithm is deterministic and is based on a simple classical primal-dual approach.

2.3 Online Prize-collecting Non-metric Facility Location

Assuming the word-of-mouth effect, the advertising company now wishes to send the ad to very few people so as every potential customer announced is not very far from one of these people. As before, some potential customers may be ignored at the cost of paying a penalty. The goal is to minimize the total costs of sending ads, penalties, *and* the total distances from the potential customers to the nearest people who have received the ad.

This scenario can be formulated as the *Online Prize-collecting Non-metric Facility Location* problem (OPC-NFL). OPC-NFL is the online prize-collecting variant of Non-metric Facility Location and is defined as follows.

Definition 4. *(OPC-NFL) Given a complete bipartite graph $G = ((F \cup D), E)$, where F is the set of facilities that may be opened and D is the set of clients arriving over time, an edge-weight function $w : E \to R^+$, a facility-opening-cost function $f : F \to R^+$, and a penalty-cost function $p : D \to R^+$. To connect client $i \in D$ to facility j, the weight $w_{(i,j)}$ of edge (i, j) is to be paid, and to open facility $j \in F$, the opening facility cost f_j is to be paid. In each step t, a client $i \in D$ arrives; OPC-NFL asks to either pay the penalty associated with i or connect i to an open facility. The goal is to minimize the total penalties, the total facility opening costs, and the total connecting costs paid.*

To the best of our knowledge, no online algorithm with non-trivial competitive ratio exists for OPC-NFL. There only is an online algorithm for the metric version, in which facilities and clients reside in a metric space and all distances respect the triangle inequality, as in (Felice et al., 2015). The latter is essential to prove the competitive ratio of the algorithm and so the result does not carry over to OPC-NFL.

OPC-NFL generalizes the *Online Set Cover* problem (OSC), due to (Alon et al., 2003), and this implies an $\Omega(\log m \log n)$ lower bound on the competitive ratio of any online randomized polynomial-time algorithm for OPC-NFL, where m is the number of facilities and n is the number of clients.

In this paper, we propose the first online algorithm for OPC-NFL, with asymptotically optimal $O(\log m \log n)$-competitive ratio, where m is the number of facilities and n is the number of clients. Our algorithm is randomized and is based on reducing OPC-NFL to the *Online Non-metric Facility Location* problem (ONFL), due to (Alon et al., 2006).

Outline. The rest of the paper is structured as follows. In Section 3, we give an overview of results related to our problems. In Sections 4, 5, and 6, we present our results for OPC-CDS, OPC-VC, and OPC-NFL, respectively. We conclude with a discussion of our results and open problems in Section 7.

3 STATE-OF-THE-ART

(Qian and Williamson, 2011) initiated the study of online *prize-collecting* Steiner problems by providing an $O(\log n)$-competitive algorithm for the *Online Prize-collecting Steiner Tree* problem (OPC-ST). (Hajiaghayi et al., 2014) proposed an online algorithm with the same competitive ratio but gave a simpler analysis. (Hajiaghayi et al., 2014) proposed a generic technique that reduces online prize-collecting Steiner problems to their corresponding fractional non-prize-collecting variants, by losing logarithmic factor in the competitive ratio. This has implied $O(\log^3 n)$-competitive and $O(\log^4 n)$-competitive randomized algorithms for the *Online Prize-collecting Node-weighted Steiner Tree* problem (OPC-NWST) and the *Online Prize-collecting Node-weighted Steiner Forest* problem (OPC-NWSF), respectively. (Hajiaghayi et al., 2013) gave a primal-dual algorithm with optimal $O(\log n)$-competitive ratio for ONWST in graphs excluding a fixed graph as minor (which include planar graphs). (Hajiaghayi et al., 2014) extended this result to OPC-NWST in graphs excluding a fixed graph as minor, for which they gave an $O(\log^2 n)$-competitive algorithm. Other well-known optimization problems were also studied in the online prize-collecting model, such as (Ausiello et al., 2008).

(Hamann et al., 2018) proposed an online randomized algorithm for a special case of OPC-CDS in which all penalties are set to infinity, the *Online Connected Dominating Set* problem (OCDS). They showed that their algorithm has asymptotically optimal $O(\log^2 n)$-competitive ratio, where n is the number of nodes. (Markarian and Kassar, 2020) later proposed an online *deterministic* algorithm for the problem with the same $O(\log^2 n)$-competitive ratio.

(JunFeng and JianHua, 2014) gave an optimal 2-approximation algorithm for the prize-collecting variant of Vertex Cover (in the offline setting). (Demange and Paschos, 2005) studied an online model of Vertex Cover, that is substantially different than the one in this paper, providing competitive ratios characterized by the maximum degree of the graph.

(Alon et al., 2006) proposed an online randomized algorithm for a special case of OPC-NFL in which all penalties are set to infinity, the *Online Non-metric Facility Location* problem (ONFL). They showed that their algorithm has $O(\log m \log n)$-competitive ratio, where m is the number of facilities and n is the number of clients. The latter is asymptotically optimal since ONFL generalizes OSC. A related problem to ONFL is the metric variant, known as the *Online Facility Location* problem (OFL), in which facilities and clients reside in a metric space and all distances respect the triangle inequality. The latter is commonly used in competitive analysis. (Meyerson, 2001) gave an $O(\log n)$-competitive randomized algorithm for OFL, where n is the number of clients. This was improved to an $O(\frac{\log n}{\log \log n})$-competitive algorithm by (Fotakis, 2008), who showed that this is the best competitive ratio achievable for the problem. Another paper of (Fotakis, 2007) gave a simple deterministic primal-dual $O(\log n)$-competitive algorithm for the problem. The prize-collecting metric variant, the *Online Prize-collecting Facility Location* problem was studied by (Felice et al., 2015), who proposed an $O(\log n)$-competitive algorithm for the problem. Their algorithm is based on previous algorithms of (Fotakis, 2008) and (Nagarajan and Williamson, 2013).

4 ONLINE PRIZE-COLLECTING CONNECTED DOMINATING SET (OPC-CDS)

In this section, we present a randomized online algorithm for OPC-CDS and analyze its competitive ratio.

4.1 Online Algorithm

The algorithm has two phases. In the first phase, we transform the given instance I into an OSC instance I' as follows.

Given an instance I of OPC-CDS containing a connected graph $G = (V,E)$, a penalty cost function $p : V \to R^+$, and a sequence of disjoint subsets of V arriving over time. We construct an instance I' of OSC as follows. The elements of I' are the nodes of V. Each node $u \in V$ is represented by two sets:

- a set containing u and all nodes adjacent to u, with cost w_u, the weight associated with u
- a set containing u, with cost p_u, the penalty associated with u

When a subset $D_t \subseteq V$ arrives at step t, the algorithm returns the set P_t containing the nodes of D_t whose penalties are paid for and the set CDS_t that contains a connected dominating set of the remaining nodes $D_t \setminus P_t$. Now, the algorithm runs the algorithm for OSC, due to (Alon et al., 2003), on I' and adds the corresponding nodes to the sets P_t and CDS_t based on the sets returned by the algorithm. Note that a node may end up covered by more than one set, meaning that its penalty might be paid for in addition to being dominated.

Note that, any OSC solution for I' of cost c is a solution of the same cost c for Phase 1 of the algorithm. Moreover, an $O(\log m \log n)$-competitive algorithm for OSC implies an $O(\log^2 n)$-competitive algorithm for Phase 1 of the algorithm, since the number of sets in I' is double the number of nodes in I ($m = 2n$).

In the second phase, we connect the dominating set nodes constructed in the first phase directly. We run the randomized algorithm for the *Online Node-weighted Steiner Tree* problem (OPC-NWST) due to (Hajiaghayi et al., 2014) on these nodes. The two phases of the algorithm are depicted below.

Online Algorithm for OPC-CDS.

Input: $G = (V, E)$ and subset $D_t \subseteq V$
Output: $P_t \cup CDS_t$

1. Run the OSC algorithm on I'. Add the nodes whose penalties are paid for to P_t and the dominating set nodes to a set S_t. If $t = 1$, assign any of the nodes in S_t as a root node r. Add all the nodes in S_t to CDS_t.
2. Run the OPC-NWST algorithm to construct a tree that connects all the nodes in S_t to r. Add all the nodes of this tree, that are not already in CDS_t, to CDS_t.

4.2 Competitive Analysis

We denote by Opt the cost of an optimal solution Opt_I for an instance I of OPC-CDS and by C_1 and C_2 the cost of the two phases of the algorithm, respectively.

Phase 1. The cost C_1 of Phase 1 of the algorithm can be bounded as follows.

$$C_1 \leq O(\log^2 n) \cdot Opt$$

Phase 2. The cost C_2 of Phase 2 of the algorithm is the cost of the Steiner tree nodes connecting the dominating set nodes constructed in the first phase. Let Opt_{St} be the cost of a minimum Steiner tree of these nodes. Since the algorithm for OPC-NWST, due to (Hajiaghayi et al., 2014), has an $O(\log^2 n)$-competitive ratio, we have that $C_2 \leq O(\log^2 n) \cdot Opt_{St}$.

It remains to compare Opt_{St} to the cost of the optimal solution Opt. The latter is not a feasible solution for Phase 2 of the algorithm. This would have been the case in the offline setting.

Remark. In the offline setting, algorithms that first find a dominating set and then run a Steiner tree algorithm to connect the dominating set, have a straightforward approximation analysis, depending on the analysis of the Steiner tree and Dominating Set algorithms themselves. This means that the approximation bounds attained, in such algorithms, depend on the approximation bounds for Dominating Set/Set Cover and Steiner Tree (see (Guha and Khuller, 1998)).

To compare Opt_{St} to Opt, we construct a Steiner tree S for the nodes of Phase 1 of the algorithm, as follows. S will contain the nodes in the optimal solution and some additional nodes. We add to S:

- all the nodes in the optimal solution
- all the nodes added in Phase 2 of the algorithm, in addition to the nodes added in Phase 1 (these are the terminals and so have weight 0 each)
- one additional node from the demand set D_t, for any t (this has weight at most w_{max})

$$\min \sum_{i \in V} x_i w_i + \sum_{e \in E} p_e y_e$$
$$\textit{Subj to:} \ \ \forall e = (i, j) \in E : x_i + x_j + y_e \geq 1$$
$$\forall i \in V, e \in E : x_i, y_e \geq 0 \ \max \sum_{e \in E} z_e$$
$$\textit{Subj to:} \ \ \forall i \in V : \sum_{e \in \gamma(i)} z_e \leq w_i$$
$$\forall e \in E : z_e \leq p_e$$
$$\forall e \in E : z_e \geq 0$$

Figure 1: LP Formulation of OPC-VC.

The cost of S is upper bounded by: $Opt + C_2 + w_{max}$. Thus, Opt_{St} is at most $Opt + C_2 + w_{max}$. Therefore, $C_2 \leq O(\log n) \cdot (Opt + C_2 + w_{max})$.

Applying asymptotic notation with simple algebra and using the fact that Opt is at least w_{min}, we conclude that

$$C_2 \leq O(\log n) \cdot \frac{w_{min}}{w_{max}}$$

By adding the two costs C_1 and C_2 of the algorithm, we conclude the following theorem.

Theorem 1. *There is an online $O(\frac{w_{max}}{w_{min}} \log^2 n)$-competitive randomized algorithm for the Online Prize-collecting Connected Dominating Set problem (OPC-CDS), where n is the number of nodes, w_{max} is the maximum node weight, and w_{min} is the minimum node weight.*

5 ONLINE PRIZE-COLLECTING VERTEX COVER (OPC-VC)

In this section, we present a deterministic online algorithm for OPC-VC and analyze its competitive ratio.

5.1 Online Algorithm

The algorithm is a classical primal-dual algorithm. The LP formulation of OPC-VC is depicted in Figure 1. x_i is the indicator variable set to 1 if node i of weight w_i belongs to the solution, and set to 0 otherwise. y_e is the indicator variable set to 1 if the penalty p_e of edge e is paid, and set to 0 otherwise. $\gamma(i)$ is the set of edges incident to node i.

The primal-dual algorithm is depicted below. Let S be the set of all edges whose penalties are paid for and all nodes that are purchased by the algorithm.

Online Algorithm for OPC-VC.

Input: $G = (V, E)$ and $e \in E$
Output: S

1. Increase the dual variable z_e of e until one of the dual constraints becomes tight.
2. Set the primal variable corresponding to each tight constraint to 1.
3. Purchase each node corresponding to a tight constraint and pay each penalty corresponding to a tight constraint.

5.2 Competitive Analysis

Let S be the primal solution constructed by the algorithm. Recall that S is the set of all edges whose penalties are paid for and all nodes that are purchased by the algorithm. We have that the dual constraint corresponding to each node $i \in S$ is tight: $w_i = \sum_{e \in \gamma(i)} z_e$. Moreover, the dual constraint corresponding to each $e \in S$ is tight: $p_e = z_e$. Thus,

$$\sum_{e \in S} p_e y_e \leq \sum_{e \in E} z_e$$

and

$$\sum_{i \in S} x_i w_i = \sum_{i \in S} \sum_{e \in \gamma(i)} z_e \leq 2 \cdot \sum_{e \in E} z_e.$$

By the Weak Duality theorem, we have that $\sum_{e \in E} z_e \leq Opt$, where Opt is the cost of the optimal solution, and hence the theorem follows.

Theorem 2. *There is an online 3-competitive deterministic algorithm for the Online Prize-collecting Vertex Cover problem (OPC-VC).*

6 ONLINE PRIZE-COLLECTING NON-METRIC FACILITY LOCATION (OPC-NFL)

In this section, we present a randomized online algorithm for OPC-NFL and analyze its competitive ratio.

6.1 Online Algorithm

Given an instance I of OPC-NFL that contains a complete bipartite graph $G = ((F \cup D), E)$, an edge-weight function $w : E \to R^+$, a facility-opening-cost function $f : F \to R^+$, and a penalty-cost function $p : D \to R^+$. The algorithm is based on transforming I into an instance I' of the *Online Non-metric Facility Locatiom* problem (ONFL), as follows.

- We add to the set F, a facility j and set its opening cost to 0.
- For each client $i \in D$ that arrives, we add an edge from i to j and set its weight to the penalty cost of i.

The algorithm is depicted below.

Online Algorithm for OPC-NFL.

Input: $G = ((F \cup D), E)$ and instance I of OPC-NFL
Output: Set of penalties, facility costs, and connecting costs paid

1. Transform I into I', as described earlier.
2. Run the algorithm for ONFL on I'.
3. Purchase all facilities and edges outputted by the ONFL algorithm. For each arriving client i, pay its associated penalty if the corresponding edge in I' is purchased by the ONFL algorithm.

6.2 Competitive Analysis

Let I be the original instance of OPC-NFL. Let Opt be an optimal solution of I and let C_{Opt} be its cost. Let I' be the new instance of ONFL generated from I as above. Let Opt' be an optimal solution of I' and let $C_{Opt'}$ be its cost.

We need to show that Opt is a feasible solution of I': Given a client i, whenever its penalty is purchased in Opt, we purchase the corresponding edge in I'; whenever a facility is opened in Opt, we open it too in I' and whenever an edge is paid for, we pay for it too in I'. This means that every time a client arrives, it is connected to at least one facility and the connecting edge is paid for by the solution Opt. Thus, Opt is a feasible solution of I'. Hence, $C_{Opt'} \leq C_{Opt}$, since every feasible solution is lower bounded by the cost of the optimal solution.

The algorithm for ONFL has an $O(\log m' \log n')$-competitive ratio, where m' is the number of facilities and n' is the number of clients. According to our reduction, $m' = m + 1$ and $n' = n$, where m is the number of facilities and n is the number of clients in the original instance I.

Let C be the cost of our solution for I. Our solution is constructed by running the algorithm for ONFL and thus $C \leq O(\log m \log n) \cdot C_{Opt'} \leq O(\log m \log n) \cdot C_{Opt}$ and the theorem below follows.

Theorem 3. *There is an online asymptotically optimal $O(\log m \log n)$-competitive randomized algorithm for the Online Prize-collecting Non-metric Facility Location problem, where m is the number of facilities and n is the number of clients.*

7 DISCUSSION AND OPEN PROBLEMS

In this paper, we have studied OPC-CDS in *general* graphs. *Connected Dominating Set* problems have been extensively studied in the offline setting. These were motivated by real-world network applications in which *geometric* graph models were used (see (Mahdian and Yan, 2011; Ambühl et al., 2006)). One research direction is to extend this study to the online setting by considering such graphs for OPC-CDS and its variants. This might yield to competitive ratios dependent on the properties of the geometric graphs rather than the number of nodes.

Another set of open problems would generate by considering other online models. We have made our study based on the *oblivious* adversary model. It is interesting to consider other weaker adversary models such as *stochastic* as in (Manshadi et al., 2010), or *random* as in (Mahdian and Yan, 2011).

Competitive analysis offers a *worst case* performance evaluation of online algorithms. It would be interesting to also consider other models. (Cheung, 2016) performed a computational study of various online algorithms for Steiner problems to reveal their performance in *average*. It would be interesting to do the same for our proposed algorithms, as in (Hamann et al., 2018)), for the *Online Connected Dominating Set* problem. Another study by (Angelopoulos, 2019) based the analysis of online algorithms on additional parameters of the problem, known as *parameterized* analysis of online algorithms. For the *Online Node-weighted Steiner Tree* problem, he showed a tight competitive ratio that depends on the maximum node weight, minimum node weight, and the number of terminals. Investigating parameterized analysis for our problems would initiate an interesting study.

REFERENCES

Alon, N., Awerbuch, B., and Azar, Y. (2003). The online set cover problem. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 100–105, New York, NY, USA. ACM.

Alon, N., Awerbuch, B., Azar, Y., Buchbinder, N., and Naor, J. S. (2006). A general approach to online network optimization problems. *ACM Trans. Algorithms*, 2(4):640–660.

Ambühl, C., Erlebach, T., Mihalák, M., and Nunkesser, M. (2006). Constant-factor approximation for minimum-weight (connected) dominating sets in unit disk graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 3–14. Springer.

Angelopoulos, S. (2019). Parameterized analysis of the online priority and node-weighted steiner tree problems. *Theory Comput. Syst.*, 63(6):1413–1447.

Ausiello, G., Bonifaci, V., and Laura, L. (2008). The online prize-collecting traveling salesman problem. *Information Processing Letters*, 107(6):199 – 204.

Cheung, S. S. (2016). *Offline and online facility location and network design*. PhD thesis, Operations Research and Information Engineering, Cornell University.

Demange, M. and Paschos, V. T. (2005). Online vertex-covering. *Theoretical Computer Science*, 332(1):83 – 108.

Felice, M. C. S., Cheung, S.-S., Lee, O., and Williamson, D. P. (2015). The online prize-collecting facility location problem. *Electronic Notes in Discrete Mathematics*, 50:151 – 156. LAGOS'15 – VIII Latin-American Algorithms, Graphs and Optimization Symposium.

Fotakis, D. (2007). A primal-dual algorithm for online non-uniform facility location. *Journal of Discrete Algorithms*, 5(1):141 – 148.

Fotakis, D. (2008). On the competitive ratio for online facility location. *Algorithmica*, 50(1):1–57.

Guha, S. and Khuller, S. (1998). Approximation algorithms for connected dominating sets. *Algorithmica*, 20(4):374–387.

Hajiaghayi, M., Liaghat, V., and Panigrahi, D. (2014). Near-optimal online algorithms for prize-collecting Steiner problems. In Esparza, J., Fraigniaud, P., Husfeldt, T., and Koutsoupias, E., editors, *Automata, Languages, and Programming*, pages 576–587, Berlin, Heidelberg. Springer Berlin Heidelberg.

Hajiaghayi, M. T., Liaghat, V., and Panigrahi, D. (2013). Online node-weighted Steiner forest and extensions via disk paintings. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 558–567.

Hamann, H., Markarian, C., Meyer auf der Heide, F., and Wahby, M. (2018). Pick, pack, & survive: Charging robots in a modern warehouse based on online connected dominating sets. In *9th International Conference on Fun with Algorithms, FUN 2018, June 13-15, 2018, La Maddalena, Italy*, pages 22:1–22:13.

JunFeng, D. and JianHua, T. (2014). A factor 2-approximation algorithm for the prize-collecting vertex cover problem. *Journal of Beijing University of Chemical Technology (Natural Science Edition)*, 41(2):120.

Korman, S. (2005). On the use of randomization in the online set cover problem. Master's thesis, Weizmann Institute of Science, Israel.

Mahdian, M. and Yan, Q. (2011). Online bipartite matching with random arrivals: An approach based on strongly factor-revealing lps. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC ?11, page 597?606, New York, NY, USA. Association for Computing Machinery.

Manshadi, V. H., Gharan, S. O., and Saberi, A. (2010). Online stochastic matching: online actions based on offline statistics. *Math. Oper. Res.*, 37:559–573.

Markarian, C. (2018). An optimal algorithm for online prize-collecting node-weighted steiner forest. In *Com-*

binatorial Algorithms - 29th International Workshop, IWOCA 2018, Singapore, July 16-19, 2018, Proceedings, pages 214–223.

Markarian, C. and Kassar, A. (2020). Online deterministic algorithms for connected dominating set & set cover leasing problems. In Parlier, G. H., Liberatore, F., and Demange, M., editors, *Proceedings of the 9th International Conference on Operations Research and Enterprise Systems, ICORES 2020, Valletta, Malta, February 22-24, 2020*, pages 121–128. SCITEPRESS.

Meyerson, A. (2001). Online facility location. In *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science*, FOCS '01, pages 426 – 431, Washington, DC, USA. IEEE Computer Society.

Nagarajan, C. and Williamson, D. P. (2013). Offline and online facility leasing. *Discrete Optimization*, 10(4):361 – 370.

Qian, J. and Williamson, D. P. (2011). An $O(\log n)$-competitive algorithm for online constrained forest problems. In *Automata, Languages and Programming - 38th International Colloquium, ICALP 2011, Zurich, Switzerland, July 4-8, 2011, Proceedings, Part I*, pages 37–48.

Sleator, D. D. and Tarjan, R. E. (1985). Amortized efficiency of list update and paging rules. *Commun. ACM*, 28(2):202–208.