An Interactive Search Game with Two Agents

Xing Gao
Department of Mathematics, Statistics, & Computer Science
University of Illinois at Chicago
Chicago, USA
xgao53@uic.edu

Lev Reyzin
Department of Mathematics, Statistics, & Computer Science
University of Illinois at Chicago
Chicago, USA
lreyzin@uic.edu

***Abstract*—We study a two-player interactive game where the players search for a target vertex in a connected undirected graph via multiple rounds of query and feedback. This two-player search game is an adaptation of the single-agent interactive learning problem [1]. Here, we analyze the game for trees under different costs: under zero-sum cost we present the one-step Nash equilibrium strategy, and show that the competition hinders both players' search progress as compared to binary search; under non-zero-sum costs, however, we present a cooperative strategy which benefits both players.**

I. Introduction

In this paper we study the problem of two players searching for a target vertex in a tree under a game-theoretic setting. Our search game is an adaption of the generalized binary search in undirected weighted graph model studied by Emamjomeh-Zadeh et al [2], which gave an algorithm for one learner/player to identify a target vertex interactively via $O(\log n)$ rounds of query and feedback, where n is the number of vertices in the graph. Specifically, in each round the player queries a vertex and receives a feedback, which either informs the player that the query vertex is the target, or reveals a neighbor of the query vertex that is on a shortest path to the target. Later Deligkas et al [3] studied the game in three new directions: 1, the response is on an approximate shortest path to target; 2, the query vertex is near optimal; 3, there are multiple targets to search for.

In this paper we extend the well-studied single-player search game to a two-player game. We are interested in how the interaction between the two players affects their query strategies, in terms of which vertices to query, and the query complexity (number of queries needed to identify the target by either player). In a competitive zero-sum game, we give the Nash equilibrium strategy and find that the query complexity is worse than that of one-player binary search. In an non-zero-sum game, we give a cooperative strategy that achieves better query complexity than one-player binary search.

This work continues several lines of research of learning target structures over graphs. One important line of previous work involves query learning, where target hypotheses are queried to an oracle, and the oracle replies with some form of feedback to the learner. This model dates back to equivalence queries [4], but also includes more modern frameworks such as correction queries [5], as well as other query types [6].

Supported in part by NSF Award CCF-1934915 and CCF-2122628.

Another perspective is from the point of view of "games on graphs" [7], where multiple players take turns making moves on a graph. Variants include coloring games [8] in which each player controls the color of a vertex, mathematical games like slither [9], or even the basic "nightclub pricing" pricing game on a path [10]. In these cases, the object is to find the Nash equilibria and to understand how changes to payoff structures impact the resulting strategies.

A. Results summary

We summarize the players' one-step Nash equilibrium strategies in terms of positions to query and corresponding probabilities in a mixed strategy, and compare progress with binary search in the table below. We only show the strategies from player x's perspective since this is a symmetric game. Refer to Section III (and IV) for the definitions of regions L, M, R on a path (and in a tree), as well as the median and quantile positions. Roughly speaking, wlog if x is to the left of y on a path, the region to the left of x is L, the region to the right of y is R, and the remaining reguin in between is M. Here we assume the middle region M is sufficiently large ($M > R$).

TABLE I: A summary of our results for 2-player interactive search on a path.

	feedback $\leftarrow$ (away from y)	feedback $\rightarrow$ (towards y)	better than binary search?
competitive (continuous 0-sum costs)	median of L	median of M w.p. $\frac{M}{M+R}$ median of R w.p. $\frac{R}{M+R}$	no
competitive (discrete 0-sum-costs)	median of L	median of M w.p. $\frac{M+R}{M+3R}$ median of R w.p. $\frac{2R}{M+3R}$	no [1]
cooperative (equal costs)	median of L	$\frac{1}{4}$-quantile of M	yes
binary search (comparator strategy)	median of L	median of $M \cup R$	N/A

[1]When $M > R$, as this table assumes, competitive search with 0-sum costs is indeed less efficient than binary search. When $M < R$, it is possible for competitive search with the discrete 0-sum cost to outperform binary search.

Our results show that a mutual non-zero-sum-cost encourages the players to make progress as a team, and their joint actions are more efficient than one-player binary search. However, under a zero-sum-cost, each player is only incentivized to outperform the other player, and even the player with better progress is slower than one-player binary search.

II. Model of the search game

Now we describe an extensive-form game where two players x and y search for a target vertex in a graph. As pointed out by Emamjomeh-Zadeh et al. [2], binary search on a path has a natural generalization to trees, and further to weighted undirected graphs with the property that the shortest path between each pair of vertices is unique, which ensures the feedback to each query is deterministic conditioned on a target location. We first consider the simplest case that the graph is a path, then generalize to trees, and discuss possible directions for arbitrary graphs.

Given an undirected graph $G = (V, E)$, we assume that first a target vertex t is chosen uniformly at random by the nature/environment, and the players' action in each round is choosing a vertex to query. Each player receives a feedback after his query, which either informs him that the query vertex is the target, or reveals a neighbor of the query vertex that is on a shortest path to the target. We assume players choose their query vertices individually and query them simultaneously. The players get **imperfect information**, in that the feedback is private to each player, but players can observe each other's actions hence deduce information about the other player's previous feedback, and adaptively choose next vertex to query based on their own feedback and observation. We assume each player begins by querying a random vertex.

Similar to the algorithm in [2] and [1], each player can define and keeps track of the weight for each vertex, which can be treated as the likelihood or probability mass of each vertex, and can be used to calculate the next vertex to query. Intuitively speaking, for a vertex v, its weight $w(v) \sim \mathbb{P}[v \text{ is the target}]$. Initially, each player defines the weight for each vertex uniformly: $\forall v \in V, w(v) = \frac{1}{n}$, where $n = |V|$. For a pair of query q and feedback z, denote $N(q, z)$ as the set of vertices inconsistent with the query-feedback pair, ie., the vertices that are eliminated from the candidate vertices for the target based on this query-feedback pair. Formally, $N(q, z) = \{v \mid z \text{ is not on a shortest path from } q \text{ to } v\}$.

At each round, the players update their weights for the vertices in the following way: suppose at the r^{th} round, player x queries vertex x_r and receives **feedback** z. He also can deduce the other player's previous feedback z' by observing the action y_r (relative to y_{r-1}). The player updates the weights by reducing the weights of the vertices that are inconsistent with his feedback and observation to 0, under the assumption that z lies on a shortest path from x_r to t, and z' lies on a shortest from y_{r-1} to t. Formally, player x sets $w(v) = 0 \; \forall v \in N(x_r, z) \cup N(y_{r-1}, z')$, and re-normalize the remaining vertices whose weights are positive, so that total weights in the graph sum up to 1.

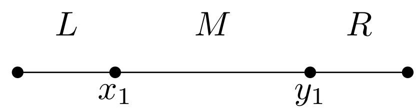

Fig. 1: Players query positions divide the path into regions L, M and R.

In a game theoretic setting, each player also incurs an implicit cost to each query (incomplete information). Naturally, the cost should reflect how far off the query vertex is from the target based on some distance measure. We will define two types of zero-sum costs which lead to different Nash equilibria in competitive games, and a non-zero-sum cost which lead to cooperative behaviors between the players. For the entirety of this paper, we are concerned with **one-step** Nash equilibria, where players are only optimizing their **immediate** reward.

III. Interactive search on a path

We begin by examining the path, as its simplicity allows us to develop intuition behind the players' strategies, yet it has straightforward generalization to more complex structures such as trees.

We use the notation illustrated below. The two players' query vertices at round r are represented by variables x_r and y_r, and we will drop the subscript when discussing an arbitrary round. Without loss of generality, we will state our results in terms of the current round being $r = 1$ and the next round as $r = 2$, and we assume that at the current round x_1 is to the "left" of y_1. These positions partition the path into regions $\{L, M, R\}$, which stand for "left," "middle," and "right" as shown in Figure 1.

For brevity we will sometimes abuse notation and use L (for example) to denote the total probability mass of region L as well, ie., $\sum_{v \in L} w(v)$. The feedback for each player either points toward the left or the right on a path, and we denote the event that player x receives a left feedback after querying position x_1 as $x_1 \leftarrow$.

In the following analysis, we abstract the path to a continuous line segment for simplicity, since every point on the line lies between two unique vertices, and we can "discretize" a position on a line to the nearest vertex. Naturally we will use a probability density function $f(\cdot)$ in place of the probability mass function $w(v)$. Without loss of generality we normalize the total length of the path at the beginning of the first round to 1 with 0 starting from the left end, so that under a uniform prior of the target location, the total probability mass of a region corresponds to its (fractional) length, so L (for example) may represent the length of region L as well.

We define the implicit cost of querying vertex q based on the squared Euclidean distance between q and the target vertex t: $d(q, t) := (q - t)^2$. We first study the game where the two players compete against each other, and we show different Nash Equilibria under two different zero-sum costs. Then we study the game where the two players cooperate under a non-zero-sum cost.

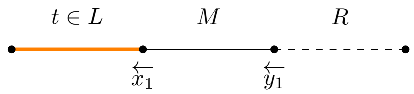

Fig. 2: $t \in L$. x has remaining region L, and y has remaining regions L and M.

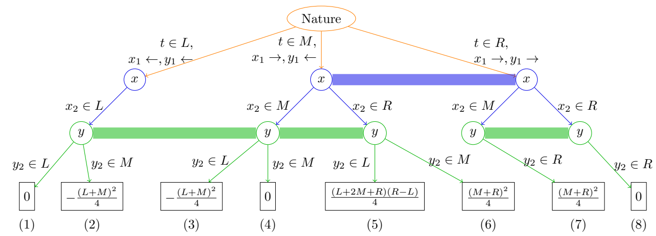

Fig. 3: The game tree of the extensive form game with zero-sum cost. Leaves represent player x's costs.

A. Competition under continuous cost

1) Continuous cost: In this section we define a continuous zero-sum cost based on the difference between each player's distance to the target. Specifically, from player x's perspective using the squared Euclidean distance, the cost is:

$$\begin{aligned} C(x \mid y, t) &= d(x,t) - d(y,t) = (x-t)^2 - (y-t)^2 \\ &= (x+y-2t)(x-y) \end{aligned} \tag{1}$$

We call this an implicit cost because even though both players' positions are public information, the cost value depends on target location which is unknown to both players. Each player's objective is to find the optimal vertex to query for the next round that minimizes the expected cost conditioned on the opponent's query position, over the randomness of the target location:

$$x_2 = \operatorname*{argmin}_x \max_y \mathbb{E}_t[C(x \mid y, t)]$$

2) Game tree: We can construct a game tree for the extensive form game based on the cost defined in the previous section. Consider an arbitrary case that the target is in region L and both players received feedback to the left in the current round. Regions M and R are eliminated for x, and region R is eliminated for y. Each player sets the probability density to 0 in the eliminated regions and re-normalizes the remaining regions (Figure 2). Suppose the regions have probability mass L and M respectively, then the posterior probability density (after query and feedback) of target location for player x is $f(t) = \frac{1}{L}$ for $t \in L$, and we have:

$$\begin{aligned} \mathbb{E}_t[C(x \mid y, t)] &= \int_L (x+y-2t)(x-y) \cdot \frac{1}{L} dt \\ &= (x-y)(x+y-L) \end{aligned} \tag{2}$$

Notice that the expected cost has a saddle point at $(\frac{1}{2}L, \frac{1}{2}L)$, which corresponds to the safety strategy $(x,y) = (\frac{1}{2}L, \frac{1}{2}L)$:

$$x = \frac{1}{2}L : \mathbb{E}_t[C(x \mid y,t)] = (\frac{1}{2}L - y)(y - \frac{1}{2}L) \leq 0 \ \forall y$$
$$y = \frac{1}{2}L : \mathbb{E}_t[-C(x \mid y,t)] = -(x - \frac{1}{2}L)(x - \frac{1}{2}L) \leq 0 \ \forall x$$

In this example we chose region L arbitrarily, which suggests in general whichever region the players query next, they would query the mid point (weighted median or $\frac{1}{2}$-quantile in general) in that region. This greatly reduces the players' actions (positions to query) to just the mid points in each of the regions. Specifically the possible query positions are $\frac{1}{2}L$, $L + \frac{1}{2}M$ and $L + M + \frac{1}{2}R$.

Plug in these query positions into Equation (2), we can construct the game tree representing the current and next round of the game (Figure 3).

For each player, the nodes following a specific feedback are in the same information set (linked by thick line), because the latest feedback is private to each player and their actions occur simultaneously, so the player cannot distinguish between these states and has to choose the same action at all nodes in the same information set.

3) Equilibrium strategy: We present a mixed strategy for both players and show that this strategy is in Nash equilibrium. Strategy for x (strategy for y is symmetric to x interchanging L and R):

If $x_1 \leftarrow$ (away from y), query median of L;
If $x_1 \rightarrow$ (towards y), query median of M with probability $\frac{M}{M+R}$, query median of R with probability $\frac{R}{M+R}$.

Theorem 1. *Both players following the above strategy in the search game under the continuous zero-sum cost defined in* (1) *is a Nash equilibrium.*

Proof. We analyze the game from x's perspective.
Case 1, $x_1 \leftarrow$:
This implies nature places target $t \in L$, and we have $y_1 \leftarrow$ as well. Based on our strategy, $x_2 = \frac{1}{2}L$; $y_2 = \frac{1}{2}L$ with probability $\frac{L}{L+M}$, and $y_2 = L + \frac{1}{2}M$ with probability $\frac{M}{L+M}$. We show that x_2 has no incentive to deviate from $\frac{1}{2}L$:

$$\mathbb{E}_{t\in L}[C(x \mid y,t)] = \mathbb{E}_y[(x-y)(x+y-L)]$$
$$= (x - \frac{L}{2})(x - \frac{L}{2}) \cdot \frac{L}{L+M} + (x - L - \frac{M}{2})(x + \frac{M}{2}) \cdot \frac{M}{L+M}$$
$$\text{Set } \frac{d}{dx} = 0 \iff \frac{1}{L+M}[2Lx - L^2 + 2Mx - LM] = 0$$
$$\implies x_2^* = \frac{L}{2} \text{ minimizes the expected cost for next round.}$$

Case 2, $x_1 \rightarrow$:
In this case x eliminates region L and considers $t \in M \cup R$. Let p_L, p_M, p_R represent the posterior probability that the target is in regions L, M, R respectively. Given $x_1 \rightarrow$ and the assumption that nature places target uniformly at random, we have $p_L = 0$, $\frac{p_M}{p_R} = \frac{M}{R}$.
We have the following game matrix based on the game tree:

	$y_2 \in L$	$y_2 \in M$	$y_2 \in R$
$x_2 \in L$	$p_L \times (1)$ $= 0$	$p_L \times (2)$ $= 0$	N/A
$x_2 \in M$ w.p. m	$p_M \times (3) =$ $\frac{-p_M \cdot (L+M)^2}{4}$	$p_M \times (4)$ $= 0$	$p_R \times (7) =$ $\frac{p_R(M+R)^2}{4}$
$x_2 \in R$ w.p. r	$p_M \times (5) =$ $\frac{p_M(L+2M+R)(R-L)}{4}$	$p_M \times (6) =$ $\frac{p_M(M+R)^2}{4}$	$p_R \times (8)$ $= 0$

Notice that row L and column L are dominated and can be ignored (greyed out in the matrix). Solve for the equilibrium strategy for x: suppose x chooses row M (query the median of M) with probability m and R (query the median of R) with probability r, based on the equalization principle, y should be indifferent between column M and column R in Nash equilibrium:

$$m \cdot 0 + r \cdot p_M \cdot \frac{(M+R)^2}{4} = m \cdot p_R \cdot \frac{(M+R)^2}{4} + r \cdot 0$$
$$\implies \frac{m}{r} = \frac{M}{R}$$

After normalization, this gives us the mixed strategy as claimed. $\square$

Notice that in case 1 when $x_1 \leftarrow$ eliminates the candidate set of vertices to a single region L without the other player, x's next action is the same as one-player binary search in region L. We argue that whenever a player's query-feedback eliminates the candidate vertices to a region that the other player is not in, the optimal next action is to do binary search in that region. Because x gained more information from his query-feedback pair in the current round and has a smaller candidate set of vertices to consider, we call him the "lead player" in this region, and binary search is information-theoretically optimal for one-player search.

B. *Competition under discrete cost*

1) Discrete cost: In this section we define a discrete zero-sum cost. A player incurs a cost of 1 if he is further to the target than his opponent, -1 if he is closer to the target, and 0 otherwise. Specifically, from player x's perspective using the squared Euclidean distance, the cost is:

$$\begin{aligned} C(x \mid y, t) &= \operatorname{sign}[(x-t)^2 - (y-t)^2] \\ &= \operatorname{sign}[(x-y)(x+y-2t)] \end{aligned} \tag{3}$$

Each player's objective is still to find the optimal vertex to query for the next round that minimizes the expected cost.

2) Game tree: We will construct a game tree similar to that of the continuous cost game. Consider an arbitrary case that the target is in region L, and both players received feedback to the left in the current round, which eliminates region M and R for x, and R for y. Suppose the regions have probability mass L and M respectively, then the posterior probability density (after query and feedback) of target location for player x is $f(t) = \frac{1}{L}$ for $t \in L$, and we have:

$$\begin{aligned} \mathbb{E}_t[C(x \mid y, t)] &= \int_L \operatorname{sign}[(x-t)^2 - (y-t)^2] \cdot \frac{1}{L} \\ &= \operatorname{sign}[(x-y)] \frac{1}{L} \int_L \operatorname{sign}\left[\frac{x+y}{2} - t\right] dt \\ &= \begin{cases} -\frac{1}{L} \cdot (x+y-L) & \text{if } x+y \le 2L \\ -1 & \text{if } x+y > 2L \end{cases} \\ &= \max\left(-\frac{1}{L} \cdot (x+y-L), -1\right) \end{aligned} \tag{4}$$

Based on current player positions and feedback, y cannot eliminate region M yet thus has no incentive to move to the left of x, which means $\operatorname{sign}[(x_2 - y_2)] \le 0$, we again have the safety strategy $(x, y) = (\frac{1}{2}L, \frac{1}{2}L)$:

$$x = \frac{1}{2}L : \mathbb{E}_t[C(x \mid y, t)] = \max\left(-\frac{1}{L} \cdot (y - \frac{1}{2}L), -1\right) \le 0 \ \forall y$$
$$y = \frac{1}{2}L : \mathbb{E}_t[-C(x \mid y, t)] = \min\left(\frac{1}{L} \cdot (x - \frac{1}{2}L), 1\right) \le 0 \ \forall x$$

The possible query positions are still $\frac{1}{2}L$, $L + \frac{1}{2}M$ and $L + M + \frac{1}{2}R$. Plug in these query positions into Equation (4), we can construct the game tree representing the current and next round of the game (Figure 4).

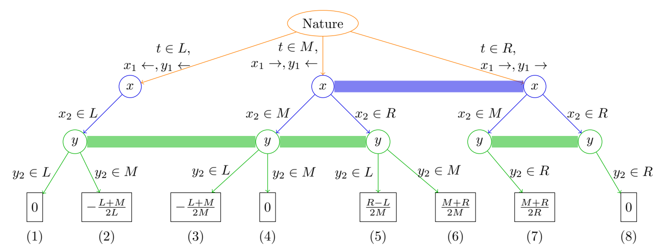

Fig. 4: The game tree of the extensive form game with zero-sum cost (discrete). Leaves represent x's costs, and all values are clipped within absolute value 1.

3) Equilibrium strategy: We present a mixed strategy for both players and claim that this strategy is in Nash equilibrium. Strategy for x (strategy for y is symmetric to x interchanging L and R):

If $x_1 \leftarrow$ (away from y), query median of L;
If $x_1 \rightarrow$ (towards y):
If $R \ge M$:

$$x_2 = \begin{cases} L + M + \frac{1}{2}R & \text{w.p. } \frac{M+R}{3M+R} \\ L + \frac{1}{2}M & \text{w.p. } \frac{2M}{3M+R} \end{cases}$$

Theorem 2. *Both players following the above strategy in the search game under the discrete zero-sum cost defined in* (3) *is a Nash equilibrium.*

Proof. We analyze the game from x's perspective.
Case 1, $x_1 \leftarrow$:
Similar to the zero-sum game under continuous cost, in this case x_2 has no incentive to deviate from $\frac{1}{2}L$.

Case 2, $x_1 \rightarrow$:
In this case x eliminates region L and considers $t \in M \cup R$. Let p_L, p_M, p_R represent the posterior probability that the target is in regions L, M, R respectively. Given $x_1 \rightarrow$ and the assumption that nature places target uniformly at random, we have $p_L = 0$, $\frac{p_M}{p_R} = \frac{M}{R}$.
We have the following game matrix based on the game tree (all values are clipped within absolute value 1):

	$y_2 \in L$	$y_2 \in M$	$y_2 \in R$
$x_2 \in L$	$p_L \times (1) = 0$	$p_L \times (2) = 0$	N/A
$x_2 \in M$ w.p. m	$p_M \times (3)$ $= -p_M \cdot \frac{L+M}{2M}$	$p_M \times (4)$ $= 0$	$p_R \times (7)$ $= p_R \cdot \frac{M+R}{2R}$
$x_2 \in R$ w.p. r	$p_M \times (5)$ $= p_M \cdot \frac{R-L}{2M}$	$p_M \times (6)$ $= p_M \cdot \frac{M+R}{2M}$	$p_R \times (8)$ $= 0$

Again, row L and column L are dominated and can be ignored (greyed out in the matrix). Solve for the equilibrium strategy for x: suppose x chooses row M (query median of M) with probability m and R (query median of R) with probability r, based on the equalization principle, y should be indifferent between column M and column R in Nash equilibrium:

$$
\begin{aligned}
&m \cdot 0 + r \cdot p_M \cdot \min\left(\frac{M+R}{2M}, 1\right) \\
&= m \cdot p_R \cdot \min\left(\frac{M+R}{2R}, 1\right) + r \cdot 0 \\
\implies \frac{r}{m} &= \begin{cases} \frac{M+R}{2M} & R \geq M \\ \frac{2R}{M+R} & R < M \end{cases}
\end{aligned}
$$

After normalization, this gives us the mixed strategy as claimed. ☐

Qualitatively speaking, the equilibrium strategies under the continuous cost and the discrete cost would both query the region that has a larger posterior probability mass with higher probability in a mixed strategy, but quantitatively speaking, the strategy under continuous cost is more "aggressive" or extreme, as it queries the larger probability region with higher probability compared to the strategy under discrete cost, and queries the smaller probability region with lower probability compared to the strategy under discrete cost.

C. Cooperation

1) Mutual cost: In this section we define a non-zero-sum cost based on the minimum distance between the players to the target, and the cost is mutual to both players. Specifically, using the squared Euclidean distance, the cost is:

$$
\begin{aligned}
C(x, y \mid t) &= \min\{d(x,t), d(y,t)\} \\
&= \min\{(x-t)^2, (y-t)^2\} \qquad (5)
\end{aligned}
$$

Given players' positions, they have the same cost value, and again this cost is implicit. The players have the same objective to find the optimal vertex to query for the next round that minimizes the expected cost conditioned on the other player's query position, over the randomness of the target location:

$$x_2 = \underset{x}{\operatorname{argmin}} \min_{y} \mathbb{E}_t[C(x, y \mid t)]$$

2) Cooperative strategy: We present a cooperative strategy for both players and show that it is better than one-player binary search.
Strategy for x (strategy for y is symmetric to x interchanging L and R, and replace $\frac{1}{4}$-quantile with $\frac{3}{4}$-quantile):

If $x_1 \leftarrow$ (away from y), query the $\frac{1}{2}$-quantile of L as the lead player in L;
If $x_1 \rightarrow$ (towards y), cooperate with y in region M at the $\frac{1}{4}$-quantile.

Here the $p-$quantile in a region M is the position $s \in M$ such that $\mathbb{P}_{t \in M}[t \leq s] = p$.

Theorem 3. *In the search game under the mutual cost defined in* (5)*, the above strategy is optimal for both players, and has lower cost than one-player binary search.*

Proof. Consider the three possibilities of target locations:
Case 1, target $t \in L$ (Figure 5):

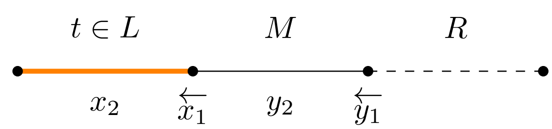

Fig. 5: Case 1, $t \in L$.

This implies $x_1 \leftarrow$ and $y_1 \leftarrow$. Player x can eliminate regions M and R, but y can only eliminate R since the players cannot communicate their current feedback. We argue that y has no incentive to query inside region L, thus his next query position should be in M: given a left feedback it's possible that $t \in M$, then y should query in M; otherwise $t \in L$, we would have $x_1 \leftarrow$ and y can trust x to act optimally as the lead player in L. In this case y's query position in M does not affect the cost for the next round, and will be determined later in Case 3. Similar to the competitive game, x doing binary search in L as a single player is optimal, and the expected cost for the next round depends on lead player x only:

$$\mathbb{E}_{t \in L}[C(x_2, y_2 \mid t)] = \int_L \left(\frac{1}{2}L - t\right)^2 \frac{1}{L} dt = \frac{1}{12}L^2$$

Case 2, target $t \in R$:
This implies $x_1 \rightarrow$ and $y_1 \rightarrow$, and this case is symmetric to Case 1: y will act as the lead player in R by querying the $\frac{1}{2}$-quantile of R, and x's next query will be in M. The expected cost for the next round depends on lead player y only:

$$\mathbb{E}_{t \in R}[C(x_2, y_2 \mid t)] = \int_R \left(\frac{1}{2}R - t\right)^2 \frac{1}{R} dt = \frac{1}{12}R^2$$

Case 3, target $t \in M$ (Figure 6):
This implies $x_1 \rightarrow$ and $y_1 \leftarrow$, and both players' next query

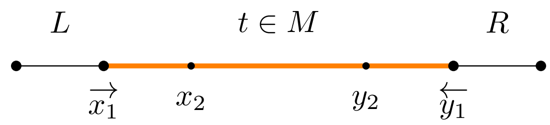

Fig. 6: Case 3, $t \in M$.

positions will be in M. To simplify calculations, we shift the position of 0 to the left end of M instead of the left end of L. Let $mid = \min\{\max\left(\frac{x+y}{2}, 0\right), M\}$, which is the mid point between x and y clipped within $[0, M]$. The expected cost depends on both players:

$$\mathbb{E}_{t\in M}[C(x,y)] = \int_{t \text{ closer to } x} d(x,t)f(t)dt + \int_{t \text{ closer to } y} d(y,t)f(t)dt$$

$$= \int_0^{mid} (x-t)^2 \frac{1}{M}dt + \int_{mid}^{M} (y-t)^2 \frac{1}{M}dt$$

$$= \frac{1}{2}x^2 - \frac{M}{4}x + \frac{1}{2}y^2 - \frac{3M}{4}y + \frac{M^2}{3}$$

Minimum attained at:

$$(x,y) = \left(\frac{1}{4}M, \frac{3}{4}M\right) \implies \mathbb{E}_{t\in M}[C(x_2,y_2)] = \frac{1}{48}M^2$$

Compare with:

$$(x,y) = \left(\frac{1}{2}M, \frac{1}{2}M\right) \implies \mathbb{E}_{t\in M}[C(x,y)] = \frac{1}{12}M^2$$

Combining the three cases gives us the cooperative strategy as claimed. Neither of the players has an incentive to deviate from this strategy, because in Case 1 and Case 2, there is always a lead player doing binary search for the next round, in a region with the smallest set of candidate vertices. The other player, even though at a disadvantageous position, shares the same low cost as the leader. In Case 3, even if both players were told that the target is in M and they perform binary search in M for the next round, the cooperative strategy achieves a lower expected cost than binary search in M individually. □

D. *Comparison of search efficiency*

We are interested in comparing the competitive strategy and the cooperative strategy with binary search under the same mutual cost defined in the previous section. We chose this mutual cost because it represents the best progress from both players for the next round, which we denote as d^*, and if we consider our game in the interactive learning framework where the two players are learners and a user/teacher is giving feedback, a lower cost in the search game corresponds to lower query complexity in interactive learning.

Again we consider the case where $t \in M$, since otherwise the lead player will be doing binary search for the next round, and this is the only case that all the strategies have distinct behaviors. Formally, we compare:

$$\begin{aligned}\mathbb{E}_{t\in M}[d^*] :=& \mathbb{E}_{t\in M}\left[\min\{(x-t)^2, (y-t)^2\}\right] \\ =& \int_0^M \min\{(x-t)^2, (y-t)^2\} \cdot \frac{1}{M}dt \\ =& \frac{1}{M}\left[\int_0^{mid} (x-t)^2 dt + \int_{mid}^{M} (y-t)^2 dt\right]\end{aligned}$$

Recall $\text{mid} = \min\{\max\left(\frac{x+y}{2}, 0\right), M\}$, the mid point between x and y clipped inside $[0, M]$. In the following sections we also assume $R = L$.

1) Cooperation:: $x = \frac{1}{4}M$, $y = \frac{3}{4}M$, $\text{mid} = \frac{1}{2}M$.

$$\mathbb{E}_{t\in M}[d^*] = \frac{1}{M}\left[\int_0^{M/2}\left(\frac{M}{4} - t\right)^2 dt + \int_{M/2}^{M}\left(\frac{3M}{4} - t\right)^2 dt\right] = \frac{M^2}{48}$$

2) Binary search:: $x = \frac{M+R}{2}$, $y = \frac{-L+M}{2}$, $\text{mid} = \frac{1}{2}M$.

$$\mathbb{E}_{t\in M}[d^*] = \frac{1}{M}\left[\int_0^{mid}\left(\frac{M-L}{2} - t\right)^2 dt + \int_{mid}^{M}\left(\frac{M+R}{2} - t\right)^2 dt\right] = \frac{M^2 + 3R^2 - 3MR}{12}$$

3) Competition:: Here we consider the competitive strategy based on the continuous zero-sum game, and compute the weighted average of the 4 possible outcomes of the players' mixed strategy:

Case 1, $x \in M$, $y \in M$ w.p. $\frac{M}{M+R} \cdot \frac{M}{L+M} = \frac{M^2}{(M+R)^2}$: $x = \frac{M}{2}$, $y = \frac{M}{2}$, $\text{mid} = \frac{M}{2}$.

$$\mathbb{E}_{t\in M}[d^*] = \int_0^M \left(\frac{M}{2} - t\right)^2 \cdot \frac{1}{M}dt = \frac{M^2}{12}$$

Case 2, $x \in R$, $y \in L$ w.p. $\frac{R}{M+R} \cdot \frac{L}{L+M} = \frac{R^2}{(M+R)^2}$: $x = M + \frac{R}{2}$, $y = -\frac{L}{2}$, $\text{mid} = \frac{M}{2}$.

$$\mathbb{E}_{t\in M}[d^*] = \frac{1}{M}\left[\int_0^{mid}\left(\frac{-L}{2} - t\right)^2 dt + \int_{mid}^{M}\left(\frac{2M+R}{2} - t\right)^2 dt\right] = \frac{M^2 + 3R^2 + 3MR}{12}$$

Case 3, $x \in M$ and $y \in L$ with probability $\frac{M}{M+R} \cdot \frac{L}{L+M} = \frac{MR}{(M+R)^2}$:

$$x = \frac{M}{2},\ y = -\frac{L}{2},\ mid = \begin{cases}\frac{M-L}{4} & M > L \\ 0 & M \le L\end{cases}.$$

$$\mathbb{E}_{t\in M}[d^*] = \frac{1}{M}\left[\int_0^{mid}\left(\frac{-L}{2} - t\right)^2 dt + \int_{mid}^{M}\left(\frac{M}{2} - t\right)^2 dt\right] = \begin{cases}\frac{M^2}{12} - \frac{L+M}{2} \cdot \frac{(M-L)^2}{16M} & M > L \\ \frac{M^2}{12} & M \le L\end{cases}$$

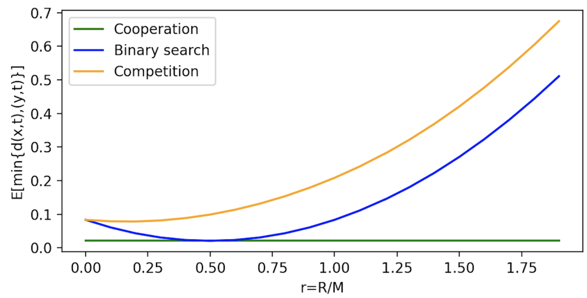

Fig. 7: Comparison between strategies.

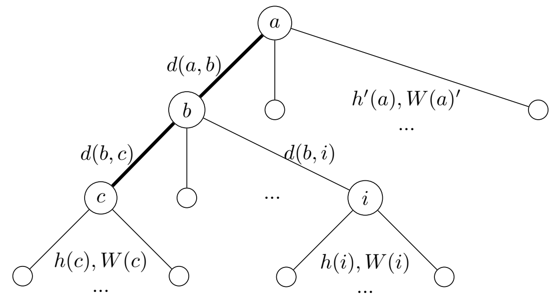

Fig. 8: Generalization to tree.

Case 4, $x \in R$ and $y \in M$ with probability $\frac{R}{M+R} \cdot \frac{M}{L+M} = \frac{MR}{(M+R)^2}$:

$$x = M + \frac{R}{2},\ y = \frac{M}{2},\ mid = \begin{cases} \frac{3M+R}{4} & M > R \\ M & M \le R \end{cases}.$$

$$\begin{aligned} &\mathbb{E}_{t\in M}[d^*] \\ =&\frac{1}{M}\left[\int_0^{mid}\left(\frac{M}{2}-t\right)^2 dt + \int_{mid}^{M}\left(\frac{2M+R}{2}-t\right)^2 dt\right] \\ =&\begin{cases} \frac{M^2}{12} - \frac{R+M}{2}\cdot\frac{(M-R)^2}{16M} & M > R \\ \frac{M^2}{12} & M \le R \end{cases} \end{aligned}$$

The overall expected cost of the mixed strategy is the sum of the 4 cases weighted by their corresponding probabilities.

4) Comparison of all three: To simplify notations suppose $R = L = rM$ and $M = 1$.

$$\text{Cooperation: } \mathbb{E}_{t\in M}\left[\min{(x-t)^2,(y-t)^2}\right] = \frac{1}{48}$$

$$\text{Binary search: } \mathbb{E}_{t\in M}\left[\min{(x-t)^2,(y-t)^2}\right] = \frac{1+3r^2-3r}{12}$$

$$\begin{aligned} &\text{Competition: } \mathbb{E}_{t\in M}\left[\min{(x-t)^2,(y-t)^2}\right] \\ &= \begin{cases} \frac{1+3r^2+3r}{12}\cdot\frac{r^2}{(r+1)^2} + \frac{1}{12}\cdot\left(1-\frac{r^2}{(r+1)^2}\right) & r \ge 1 \\ \frac{1+3r^2+3r}{12}\cdot\frac{r^2}{(r+1)^2} + \frac{1}{12}\cdot\frac{1}{(r+1)^2} \\ \quad + \frac{8-3(r+1)(r-1)^2}{96}\cdot\frac{2r}{(r+1)^2} & r < 1 \end{cases} \end{aligned}$$

In summary, in terms of the distance from the closest player to the target, as we can see from Figure 7 over various values of $r = \frac{R}{M}$, the cooperative strategy is better than binary search, and binary search is better than the competitive strategy.

IV. Interactive Search on Trees

A. Reduction from trees to a path

In the case that the graph is a tree, we can reduce the game on a tree to the game on a path naturally, since there is a unique shortest path between any pair of vertices in a tree. The regions are well-defined: L is the subtree rooted at x excluding the branch containing y, R is the subtree rooted at y excluding the branch containing x, and the rest of the graph is M.

In a subtree T, the weighted median is defined as the vertex with minimum weighted distance to all the other vertices: $q := \text{argmin}_{v\in T}\sum_{u\in T} d(v,u)\cdot w(u)$, which corresponds to the $\frac{1}{2}$-quantile if the weights of vertices represent probability masses. A path is a special case of a tree, and the players' strategies on a path generalize naturally to a tree: to query the median of region L or R, each player would query the median in the respective subtree. In the region M, we can reduce the entire region to the shortest path between x and y, and each vertex v on the path would represent the subtree rooted at v, ie., "condense" the total probability mass of the subtree rooted at v to the vertex v on the path. The median or $\frac{1}{4}, \frac{3}{4}$-quantile positions can be found by recursively expanding condensed vertex into a subtree.

B. Finding the median in a subtree

Now we discuss the details on finding the median (or quantile) vertex in a subtree. Recall that $w(v)$ is the weight (probability mass) of vertex v. Let $T(a)$ denote the subtree rooted at a, and let $W(a)$ be the total probability mass of vertices in $T(a)$, ie. $W(a) := \sum_{v\in T(a)} w(v)$. Let the height $h(a)$ be the average distance from a to vertices in $T(a)$:

$$\begin{aligned} h(a) :&= \mathbb{E}_{v\in T(a)}[d(a,v)] = \mathbb{E}[d(a,v)|v \in T(a)] \\ &= \sum_{v\in T(a)} d(a,v)\cdot\frac{w(v)}{W(a)} \end{aligned}$$

Since the distance from a to a node $c \in T(b)$ where b is a child node of a is $d(a,c) = d(a,b)+d(b,c)$ (Figure 8), we can also express the height $h(a)$ recursively based on the heights of vertex a's child nodes (Children(a)):

$$h(a) = \frac{1}{W(a)}\left[\sum_{b\in\text{Children}(a)} [h(b)+d(b,a)]\cdot W(b)\right]$$

The median in subtree T can be found recursively:
Consider subtree $T = T(a)$ with total weight $W = W(a)$, and a chain of vertices a, b, c where $b \in \text{Children}(a)$ and $c \in \text{Children}(b)$. Let $h'(a)$ and $W'(a)$ represent the height and weight of subtree rooted at a ignoring the subtree $T(b)$. See

Figure 8. We compare $h(b)$ and $h(c)$:

$$\begin{aligned}
&\mathbb{E}[d(b,t)|t\in T]\\
&=\Big(\sum_{i\in \text{Children}(b)}[h(i)+d(i,b)]\frac{W(i)}{W}\Big)\\
&\qquad\qquad + d(b,b)\frac{w(b)}{W}+[h'(a)+d(a,b)]\frac{W'(a)}{W}\\
&=[h(c)+d(c,b)]\frac{W(c)}{W}\\
&+\Big(\sum_{i\in \text{Children}(b)\setminus c}[h(i)+d(i,b)]\frac{W(i)}{W}\Big)+[h'(a)+d(a,b)]\frac{W'(a)}{W}
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[d(c,t)|t\in T]\\
&=h(c)\frac{W(c)}{W}+\Big(\sum_{i\in \text{Children}(b)\setminus c}[h(i)+d(i,b)+d(b,c)]\frac{W(i)}{W}\Big)\\
&\qquad + d(b,c)\frac{w(b)}{W}+[h'(a)+d(a,b)+d(b,c)]\frac{W'(a)}{W}\\
&=\mathbb{E}[d(b,t)|t\in T]-d(b,c)\frac{W(c)}{W}\\
&+\Big(\sum_{i\in \text{Children}(b)\setminus c}d(b,c)\frac{W(i)}{W}\Big)+d(b,c)\frac{w(b)}{W}+d(b,c)\frac{W'(a)}{W}\\
&=\mathbb{E}[d(b,t)|t\in T]-d(b,c)\Big(\frac{2W(c)}{W}-1\Big)
\end{aligned}$$

If $W(c)\geq\frac{1}{2}W$, $h(c)=\mathbb{E}[d(c,t)|t\in T]\leq\mathbb{E}[d(b,t)|t\in T]=h(b)$, which means node c is a better candidate for the median node than b. Note that the median only depends on the weights of vertices, not the lengths of edges.

To recursively find the weighted median in $T(a)$, start with root $b=a$, if there exists $c\in$ Children(b) st. $W(c)\geq\frac{1}{2}W(a)$, set $b=c$ and continue with recursion down the subtree. Otherwise return b as the weighted median, which is the lowest vertex with $W(b)\geq\frac{1}{2}W(a)$ (the lowest vertex whose weight below it is heavier than $\frac{1}{2}$ of T).

Using a similar idea, to find the $\frac{1}{4}$-quantile vertex in M from player x's perspective, start with root of subtree x and set $b=x$, if there exists $c\in$ Children(b) st. $W(c)\geq\frac{3}{4}M$, set $b=c$ and continue with recursion. Otherwise return b as the $\frac{1}{4}$-quantile of M, which is the lowest vertex b with $W(b)\geq\frac{3}{4}M$.

V. Future directions

A. General graph

In this paper we studied trees. In general graphs, the regions L, M and R might be ambiguous or have overlap due to the existence of cycles. One idea is to construct a probabilistic tree in the following way:

- For a query-feedback pair (q,z), let $S(q,z)$ be the set of vertices consistent with the query-feedback pair:

$$S(q,z)=V\setminus N(q,z)=\{v\mid z\text{ is on a shortest path from }q\text{ to }v\}$$

.

- Let $\tilde{L}=\bigcup_{(x,v)\in E:y\notin S(x,v)}S(x,v)$, which are the vertices consistent with a feedback pointing away from y from x's perspective. Similarly let $\tilde{R}=\bigcup_{(y,u)\in E:x\notin S(y,u)}S(y,u)$. Let $\tilde{M}=\tilde{L}\cap\tilde{R}$, the "ambiguous" region.
- For each neighbor v of x, let $m=\frac{w\left(S(x,v)\cap\tilde{M}\right)}{w(S(x,v))}$ (the fractional weight of $S(x,v)$ in $\tilde{M}$), and place $S(x,v)$ as a subtree rooted at v in region L with weight multiplied by $1-m$, and place $S(x,v)$ as a subtree rooted at v in region M with weight multiplied by m.
- If some vertex u appears in $S(x,v_j)$ for multiple neighbors v_j of x, split the weight of u evenly among all $S(x,v_j)$ containing it.

To analyze the correctness and how the strategies generalize to general graphs based on this idea remains future work.

B. Noisy feedback

We studied the interaction between two players under the assumption that the feedback is truthful. Previous works [1], [2], [11] considered feedback which is correct with probability $p>\frac{1}{2}$ and adversarially incorrect with probability $1-p$. To generalize our work to the noisy setting, players can update the weights of vertices in the following way: multiply the weights of vertices inconsistent with the feedback by $(1-p)$ instead of setting them to 0, and multiply the weights of the vertices consistent with the feedback by p instead of keeping them the same. Finally re-normalize the weights. A complete analysis of the generalized strategies under the noisy setting remains future work.

References

[1] E. Emamjomeh-Zadeh and D. Kempe, "A general framework for robust interactive learning," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 7085–7094.

[2] E. Emamjomeh-Zadeh, D. Kempe, and V. Singhal, "Deterministic and probabilistic binary search in graphs," in *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, D. Wichs and Y. Mansour, Eds. ACM, 2016, pp. 519–532.

[3] A. Deligkas, G. B. Mertzios, and P. G. Spirakis, "Binary search in graphs revisited," *Algorithmica*, vol. 81, no. 5, pp. 1757–1780, 2019.

[4] D. Angluin, "Learning regular sets from queries and counterexamples," *Inf. Comput.*, vol. 75, no. 2, pp. 87–106, 1987.

[5] L. Becerra-Bonache, A. Dediu, and C. Tîrnăucă, "Learning DFA from correction and equivalence queries," in *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, ser. Lecture Notes in Computer Science, Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, and E. Tomita, Eds., vol. 4201. Springer, 2006, pp. 281–292.

[6] L. Reyzin and N. Srivastava, "Learning and verifying graphs using queries with a focus on edge counting," in *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 1-4, 2007, Proceedings*, ser. Lecture Notes in Computer Science, M. Hutter, R. A. Servedio, and E. Takimoto, Eds., vol. 4754. Springer, 2007, pp. 285–297.

[7] A. E. Roth, "Two-person games on graphs," *Journal of Combinatorial Theory, Series B*, vol. 24, no. 2, pp. 238–241, 1978.

[8] K. Chaudhuri, F. C. Graham, and M. S. Jamall, "A network coloring game," in *Internet and Network Economics, 4th International Workshop, WINE 2008, Shanghai, China, December 17-20, 2008. Proceedings*, ser. Lecture Notes in Computer Science, C. H. Papadimitriou and S. Zhang, Eds., vol. 5385. Springer, 2008, pp. 522–530.

[9] W. N. Anderson, “Maximum matching and the game of slither,” *Journal of Combinatorial Theory, Series B*, vol. 17, no. 3, pp. 234–239, 1974. [Online]. Available: https://www.sciencedirect.com/science/article/pii/009589567490029X

[10] A. R. Karlin and Y. Peres, *Game theory, alive*. American Mathematical Soc., 2017, vol. 101.

[11] E. Emamjomeh-Zadeh, D. Kempe, M. Mahdian, and R. E. Schapire, “Interactive learning of a dynamic structure,” in *Algorithmic Learning Theory, ALT 2020, 8-11 February 2020, San Diego, CA, USA*, ser. Proceedings of Machine Learning Research, A. Kontorovich and G. Neu, Eds., vol. 117. PMLR, 2020, pp. 277–296. [Online]. Available: http://proceedings.mlr.press/v117/emamjomeh-zadeh20a.html