

Catching the Drunk Robber on a Graph

Natasha Komarov

Department of Mathematics
Dartmouth College
Hanover, NH 03755-3551 USA

natalie.komarov.gr@dartmouth.edu

Peter Winkler*

Department of Mathematics
Dartmouth College
Hanover, NH 03755-3551 USA

peter.winkler@dartmouth.edu

Submitted: May 21, 2013; Accepted: Aug 12, 2014; Published: Aug 21, 2014

Mathematics Subject Classifications: 05C57, 05C81, 91A43

Abstract

We show that the expected time for a smart “cop” to catch a drunk “robber” on an n -vertex graph is at most $n+o(n)$. More precisely, let G be a simple, connected, undirected graph with distinguished points u and v among its n vertices. A cop begins at u and a robber at v ; they move alternately from vertex to adjacent vertex. The robber moves randomly, according to a simple random walk on G ; the cop sees all and moves as she wishes, with the object of “capturing” the robber—that is, occupying the same vertex—in least expected time. We show that the cop succeeds in expected time no more than $n+o(n)$. Since there are graphs in which capture time is at least $n-o(n)$, this is roughly best possible. We note also that no function of the diameter can be a bound on capture time.

Keywords: Cops and Robbers; pursuit games on graphs; random walks

1 Introduction

The game of cops and robbers on graphs was introduced independently by Quilliot [10] and Nowakowski and Winkler [8], and has generated a great deal of study; see, e.g., [15, 16, 17, 18]. In the original formulation a cop and robber move alternately and deliberately, with full information, from vertex to adjacent vertex on a graph G , with the cop trying to capture the robber and the robber trying to elude the cop. In this work, all graphs are assumed to be connected, simple (no loops or multiple edges) and undirected. A graph is said to be “cop-win” if there is a vertex u such that for every v , the cop beginning at u can capture the robber beginning at v .

*Research supported by NSF grant DMS-0901475

In addition to their obvious role in pursuit games, cop-win graphs (which are also known as “dismantlable” graphs) have appeared in diverse places including statistical physics [3]. In the present work, we consider a variation suggested [7] by Ross Churchley of the University of Victoria, in which the robber is no longer in control of his fate; instead, at each step he moves to a neighboring vertex chosen uniformly at random. We may therefore imagine that the robber is in fact a drunk—one who is too far gone to have an objective.

On any graph, the drunk will be caught with probability one, even by a cop who oscillates on an edge, or moves about randomly; indeed, by any cop who isn’t actively trying to lose. The only issue is: how long does it take? The lazy cop will win in expected time at most $4n^3/27$ (plus lower-order terms), since that is the maximum possible expected hitting time for a random walk on an n -vertex graph [2]; the same bound applies to the random cop [4]. It is easy to see that the greedy cop who merely moves toward the drunk at every step can achieve $O(n^2)$; in fact, we will show that the greedy cop cannot in general do better. Our smart cop, however, gets her man in expected time $n + o(n)$. Note that when the adversaries play on a lollipop graph consisting of a clique of size $cn^{1/3}$ (for some constant $c \in \mathbb{R}$) with a path of length $n - cn^{1/3}$ attached at one end, with the drunk starting in the clique and the cop starting at the opposite endpoint of the path, the expected capture time will be $n - \Theta(n^{1/3}) = n - o(n)$, and we conjecture that this is worst possible.

2 Preliminaries

In this variation, a “move” (as in chess) will consist of a step by the cop followed by a (uniformly random) step by the drunk. Capture or “arrest” takes place when the cop lands on the drunk’s vertex or vice-versa, and the capture time T is the number of the move at which this takes place.

Let us consider some examples. (1) Suppose G is the path P_n on n vertices, with u and v its endpoints. Then the cop will (using any of the algorithms we consider later) move along the path until she reaches the drunk; this will take expected time about $n - \sqrt{n}$ since a random walk on a path will on average progress about distance \sqrt{t} in time t .

(2) Let G be the complete balanced bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$, with the cop and the drunk beginning on the same side. Then the poor cop will find herself always moving to the opposite side from her quarry until, finally, he accidentally runs into her; since the latter event occurs with probability about $2/n$, arrest takes on average $n/2$ steps.

The reader may feel with some justification that we are being unrealistic in not allowing the cop to stay put; in example (2), sitting for one move would enable her to catch the drunk on the next move. Ultimately, we force the cop to move at each step in order to hold her to the same constraints as her quarry’s, and because it gives us the strongest results. Our bounds still apply when the cop, the drunk, or both are allowed to stay put on any move.

Even when the cop is permitted to idle, she cannot expect to catch the drunk in time bounded by a function of the diameter of G . Example (3), let G be the incidence graph

of a projective plane of order n . A projective plane P of order n is a collection of objects called “points” and sets of points called “lines” satisfying the following conditions:

1. Two points determine a unique line.
2. Two lines intersect in a unique point.
3. Every line consists of exactly $n + 1$ distinct points.
4. Every point lies on exactly $n + 1$ distinct lines.

Furthermore [5],

5. P contains exactly $n^2 + n + 1$ distinct points.
6. P contains exactly $n^2 + n + 1$ distinct lines.

Projective planes of order n are known to exist for $n = p^a$ for any prime number p and positive integer a [13]. The incidence graph G of P is therefore a graph with $2(n^2 + n + 1)$ vertices, with adjacency relation $u \sim v$ if u is a point in P and v is a line that goes through u , or vice versa. Such graphs have bounded diameter but unbounded expected capture time:

Claim 1. $\text{diam}(G) = 3$.

Proof. Let a, b be two points in P . By condition (1) above, a and b both lie on a common line, so $d(a, b) = 2$. If a, b are instead two lines in P , then condition (2) says that a and b intersect at a common point. Finally, if a is a point and b is a line in P , then either a lies on b and so $d(a, b) = 1$ or there is another point, c , which does lie on b . But by the previous argument, $d(a, c) = 2$ and so $d(a, b) = 3$. \square

Claim 2. The girth of G is 6.

Proof. Note that G has no odd cycles by the independence of the set of points (and respectively, set of lines). Now assume for sake of contradiction that G contains a cycle of length 4. Then there are two points p_1, p_2 and two lines ℓ_1, ℓ_2 such that $p_1, \ell_1, p_2, \ell_2, p_1$ forms a cycle. But this contradicts condition (2) since ℓ_1 and ℓ_2 must intersect in p_1 as well as p_2 . \square

Claim 3. G is regular of degree $r = n + 1 \approx \sqrt{|V(G)|/2}$.

Proof. By conditions (3) and (4). \square

Claim 4. The expected capture time on G is at least r .

Proof. When the cop gets to distance 2 of her quarry, the drunk has only one bad move out of r ; the rest keep him at distance at least 2. (Similarly, if the cop gets to distance 1, bypassing ever being at distance 2, the drunk still has only one bad move out of r , the rest of which keep him at distance 1.) Hence the cop’s expected capture time cannot be any lower than r (the expected number of independent Bernoulli trials, each with success probability $1/r$, until success is achieved). \square

It is not hard to verify that on any regular graph, the greedy cop—who minimizes her distance to the drunk at each move—wins in expected time at most linear in n . If G is regular of degree r , its diameter cannot exceed $\frac{3n-r-3}{r+1}$ [11]. Since the drunk will step toward the cop with probability at least $1/r$ at each move, resulting (after her response) in a decrease of 2 in their distance, the expected capture time is bounded by $r \cdot \text{diam}(G)/2 < 3n/2$.

The linear bound also holds on trees. Suppose that the cop begins at c_0 and the drunk at r_0 . Since every pair of points are joined by a unique path on a tree, the drunk will never access the vertices v for which c_0 is on the unique r_0 — v path. Therefore, the cop does not need to revisit a vertex in her attempt to capture the drunk in minimal time, and the game takes time at most $\text{diam}(G)$.

For general G , one can guarantee only that at a given point in time the drunk will step toward the cop with probability at least $1/\Delta$, giving a bound of order n^2 for the greedy cop. That may appear to be a gross overestimate, especially in light of the special cases discussed above, but a graph with many high-degree vertices can still have large diameter. For example, consider the following graph.

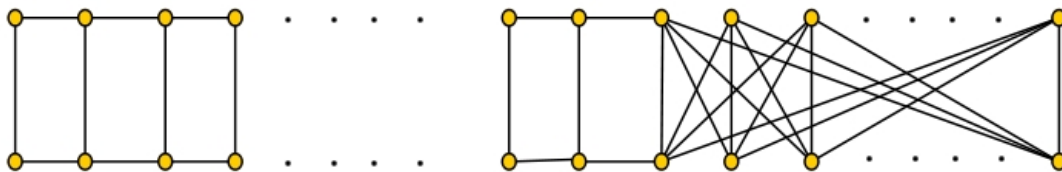


Figure 1: The Ladder to the Basement

The “ladder” in this graph consists of two copies of the path $P_{n/4}$ with each pair of corresponding vertices connected by an edge. The “basement” consists of a complete bipartite graph, $K_{\lfloor n/4 \rfloor, \lfloor n/4 \rfloor}$. We begin with the drunk inside the basement, and the cop on the far end of the ladder. While the drunk is meandering inside the basement, the cop—staying true to her goal of minimizing the distance between her and the drunk at each step—is alternating between the two paths. Note that we assume she makes the foolish choice when she is presented with several options by her algorithm. It takes the drunk $n/4$ moves on average to leave the basement, and each time this occurs, the cop

will decrease the distance by 2 by traveling along her current path. Therefore the capture will require an average of about $(n/4)^2/2$ steps.

3 A Smarter Cop

3.1 Intuition

As noted in the example of the “ladder to the basement” graph in Section 2, a foolish greedy cop can be foiled by her desire to “retarget” too often. That is, since she updates the target vertex (to which she is trying to minimize her distance) at each step, she is made indecisive by an indecisive drunk. One natural solution to this problem would be to walk directly toward the robber’s initial position in the basement for several steps before retargeting. Changing the frequency of her retargeting in this way, the cop makes steady progress, ultimately catching the drunk in time at most $n+o(n)$.

In general, if a cop and drunk begin at distance d on a graph, and the cop proceeds by retargeting every four steps, then by Lemma 3.4 below, it would take at most $4(4n^{2/3})(d-3)$ moves to get down to distance less than four. Since d can be as large as $n-1$, this bound does not suffice to yield our promised bound on expected capture time of $n+o(n)$. We have the cop do something else to get her distance to the drunk down without spending too much time doing so—hence the following four-stage strategy.

For $i \in [4]$, let T_i be the time spent in Stage i and D_i be the distance between the two players at the end of Stage i . In the first stage, the cop heads directly for the drunk’s initial position, x , so that $T_1 \leq \text{diam}(G)$. In the meantime, the drunk has gone somewhere else, and so suppose that by the time that the cop reaches x , the drunk is at y . Now we are in Stage 2, and the cop heads for y . We show $\mathbb{E}[T_2] = o(n)$. During Stage 3, the cop updates her “target” every four steps, and we show that the expected time for this stage, $\mathbb{E}[T_3]$, is again bounded by $o(n)$. This stage ends when we are at distance at most three from the drunk. In Stage 4, the cop waits for the drunk to make an error, which happens in expected time at most Δ and results in the capture of the drunk. All together, this cop captures the drunk in expected time $n + o(n)$. We will refer to the progress made by the cop in the first two stages as “gross progress,” and in the last two stages as “fine progress.” In order to prove the bounds claimed above, it will be beneficial to have a few lemmas.

3.2 Gross Progress

Suppose that the drunk starts on vertex u and the cop starts at v . As noted in the set-up of the previous section, in the first stage of the cop’s strategy, she is concerned only with getting to u (even if this may not decrease her distance from the drunk at the end of the stage). Clearly the time this takes is equal to $T_1 = d(v, u) \leq \text{diam}(G)$. We would like to get a bound on $\mathbb{E}[D_1]$, the expected distance between the cop and the drunk at the end of this stage. For that, the following lemma will prove quite useful.

Lemma 3.1. *Let $T_{n,t}$ be the distance covered in time t by a random walk on a (connected) graph with n vertices. Then $\mathbb{E}[T_{n,t}] < 1 + \sqrt{t}\sqrt{1 + 5 \log n}$.*

Proof. Let $p^t(x, y)$ be the probability that a random walk that starts at vertex x will be at vertex y in exactly t steps. The Varopoulos-Carne bound [12], as formulated in [9], says

$$p^t(x, y) \leq \sqrt{e} \sqrt{\frac{\deg(y)}{\deg(x)}} \exp\left(-\frac{d(x, y)^2}{2t}\right)$$

where $d(x, y)$ is the graph distance between the two vertices. Therefore, if we consider the random walk x_0, x_1, \dots, x_t on a graph of size n and let $c \in \mathbb{R}$ be any constant, we have the following bound as a corollary of Varopoulos-Carne:

$$\begin{aligned} \mathbb{P}(d(x_0, x_t) \geq c\sqrt{t}) &= \sum_{y:d(x_0, y) \geq c\sqrt{t}} p^t(x_0, y) \\ &\leq \sum_{y:d(x_0, y) \geq c\sqrt{t}} \sqrt{e} \sqrt{\frac{\deg(y)}{\deg(x_0)}} \exp\left(-\frac{d(x_0, y)^2}{2t}\right) \\ &< \sum_{y:d(x_0, y) \geq c\sqrt{t}} \sqrt{e} \sqrt{n} \exp\left(-\frac{c^2 t}{2t}\right) \\ &< n^{3/2} \exp\left(\frac{1 - c^2}{2}\right) \end{aligned}$$

Letting $c = \sqrt{1 + 5 \log n}$ therefore yields that $\mathbb{P}(d(x_0, x_t) \geq \sqrt{1 + 5 \log n} \sqrt{t}) < \frac{1}{n}$.

Note that $\mathbb{E}[d(x_0, x_t)] \leq pn + (1 - p)c\sqrt{t}$, where $p = \mathbb{P}(d(x_0, x_t) \geq c\sqrt{t})$, so we have

$$\begin{aligned} \mathbb{E}[d(x_0, x_t)] &\leq \frac{1}{n}n + c\sqrt{t} \\ &= 1 + \sqrt{t}\sqrt{1 + 5 \log n} \end{aligned}$$

as desired. □

This bound is not tight, but it will be good enough to give us the $o(n)$ bound we seek on $\mathbb{E}[T_2]$.

Recall that D_1 is the distance between the two players at the end of Stage 1. Note that this is equivalent to the distance between the drunk's initial position and his position at the end of Stage 1. We have the following immediate corollary of Lemma 3.1.

Corollary 3.2. $\mathbb{E}[D_1] \leq 1 + \sqrt{n}\sqrt{1 + 5 \log n}$.

Now the cop enters Stage 2. We would like to bound $\mathbb{E}[D_2]$. Note that this is equivalent to the expected distance traveled by the drunk in Stage 2.

Corollary 3.3. $\mathbb{E}[D_2] < (5 \log n)^{3/4} n^{1/4}$

Proof. Using Lemma 3.1, Jensen’s inequality for concave functions, and Corollary 3.2, we get

$$\begin{aligned} \mathbb{E}[D_2] &\leq \sum_{k=0}^n \mathbb{P}(D_1 = k)(1 + \sqrt{k}\sqrt{1 + 5 \log n}) \\ &= 1 + \sqrt{1 + 5 \log n} \mathbb{E}[\sqrt{D_1}] \\ &\leq 1 + \sqrt{1 + 5 \log n} \sqrt{\mathbb{E}[D_1]} \\ &\leq 1 + \sqrt{1 + 5 \log n} \sqrt{1 + \sqrt{n}\sqrt{1 + 5 \log n}} \\ &< (5 \log n)^{3/4} n^{1/4} \end{aligned}$$

□

Now we are done with the “gross progress” that the cop makes in Stages 1 and 2. Note that the total expected time to complete these two stages is bounded by

$$\mathbb{E}[T_1] + \mathbb{E}[T_2] \leq \text{diam}(G) + 1 + \sqrt{n}\sqrt{1 + 5 \log n}.$$

3.3 Fine Progress

At the conclusion of stage 2, the cop’s approach changes. Now she **retargets** every 4 moves. We make this notion precise in the following manner.

Let j be the time that has elapsed from the beginning of her search. For each integer $j \geq 1$ let x_j, y_j be the drunk’s and cop’s positions, respectively, at time j (with it being the cop’s turn to move). Then in Stage 3, while $d(x_j, y_{j-1}) \geq 4$, for all j of the form $4i + 1$ for some $i \geq 0$, the cop chooses x_{4i+1} as her **target** and proceeds along a geodesic toward that target for the next four steps. Consequently, the cop’s target changes every 4 moves, so that for each integer $i \geq 0$, she has target x_{4i+1} at times $4i + 1, 4i + 2, 4i + 3$, and $4i + 4$. If at time $j = 4i + 1$, $d(x_j, y_{j-1}) < 4$, Stage 3 terminates and the cop’s strategy moves into Stage 4, which will be described after the following lemma.

Lemma 3.4. *Let G be any graph and let $x_0 \in V(G)$ be any vertex in G . Let $\{x_0, x_1, x_2, \dots\}$ be any random walk on G beginning at x_0 . Then $\mathbb{P}(d(x_0, x_4) < 4) \geq 1/s$, where $s = 4n^{2/3}$.*

Before we prove this lemma, note that we could not get away with looking at the first three steps of a random walk. That is, we could not get a useful bound for $\mathbb{P}(d(x_0, x_3) < 3)$. Consider the following example: we have a graph G with a vertex x_0 . Let A_k be the set of vertices at distance k from x_0 . Suppose that G looks like Figure 2. That is in G , $|A_1| = 1$ and $|A_2| = |A_3| = \frac{n-2}{2}$. Call any step by the random walker that guarantees $d(x_0, x_3) < 3$ a “stall.” Then the probability of a stall occurring at the second step is $\frac{1}{(n-2)/2+1} = \frac{2}{n}$, and the probability of a stall occurring at the third step

is $\left(1 - \frac{2}{n}\right) \left(\frac{2}{n}\right)$ since for each vertex in A_2 and A_3 , there is one edge on the path toward x_0 and $\frac{n-2}{2}$ edges leading farther away from x_0 . So then $\mathbb{P}(d(x_0, x_3) < 3) = \frac{2}{n} + \left(1 - \frac{2}{n}\right) \frac{2}{n} < \frac{4}{n}$.

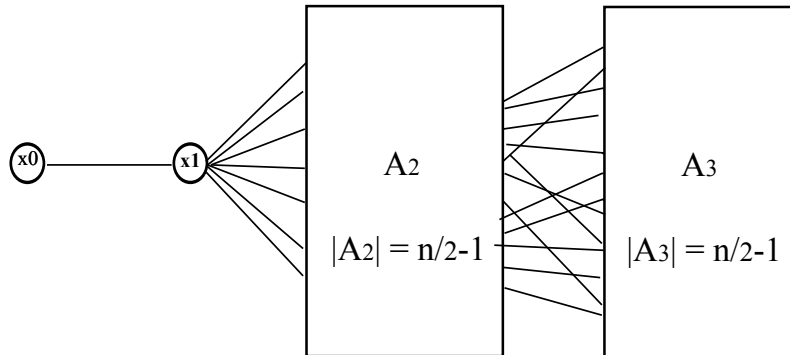


Figure 2: At least 4 steps are required to secure a useful bound on a random walk’s progress.

We now return to the proof of Lemma 3.4.

Proof. We proceed by assuming a graph G and a vertex $x_0 \in V(G)$ are such that there is a random walk $\{x_0, x_1, \dots\}$ with the property $\mathbb{P}(d(x_0, x_4) < 4) < 1/s$, and we shall derive a contradiction.

Let A_k be the set of vertices at distance k from x_0 , and let $a_k = |A_k|$ for all k . We adapt the terms **in-degree** and **out-degree** to mean the following:

Let $v \in A_k$. Then the in-degree of v is $\deg^-(v) = |N_G(v) \cap A_{k-1}|$ and the out-degree of v is $\deg^+(v) = |N_G(v) \cap A_{k+1}|$. We will use the notation p_G for the quantity under investigation, $\mathbb{P}(d(x_0, x_4) < 4)$, and for a vertex $v \in V(G)$, we define $p_k(v)$ to be the quantity $\mathbb{P}(d(x_0, x_4) < 4 | x_k = v)$. Note that $p_0(x_0) = p_G$ and $p_k(v) = 1$ if $v \in A_j$ for some $j < k$. Finally, we call any step by the random walker that guarantees $d(x_0, x_4) < 4$ a “stall.”

We will break this proof into several statements.

Claim 1. *Let G' be the graph defined by removing all edges between x_0 and all but one vertex, x_1 , where $p_1(x_1) = \min_{v \in A_1} p_1(v)$. Then $p_{G'} \leq p_G$.*

Proof. Since $p_G < 1/s$, there must exist a vertex $v \in A_1$ with $p_1(v) < 1/s$. Choose x_1 such that $p_1(x_1) = \min_{v \in A_1} p_1(v)$ and define G' as in the statement of the claim. Note that

$$p_{G'} = p_1(x_1) \leq \frac{1}{a_1} \sum_{v \in A_1 \subseteq V(G)} p_1(v) = p_G. \quad \square$$

Claim 2. Let G'' be the induced subgraph of G' with $V(G'') = V(G') - \bigcup_{k>4} A_k$ and with all edges removed except those that are between a vertex in A_{k-1} and a vertex in A_k for $k \in [4]$. Then $p_{G''} \leq p_{G'}$.

Proof. Let \hat{G}' be the induced subgraph of G' on the vertices $V(G') - \bigcup_{k>4} A_k$.

Then since $\mathbb{P}(x_t \in A_k) = 0$ when $t \leq 4$ and $k \geq 5$ (so in particular, $p_{\hat{G}'}$ and $p_{G'}$ depend only on the first four steps of a random walk originating at x_0), we have that $p_{\hat{G}'} = p_{G'}$.

Let $k \in [4]$ and let $v \in A_k$ be a vertex in $V(\hat{G}')$ with $N_{\hat{G}'}(v) \cap A_k \neq \emptyset$. If no such vertex exists then $\hat{G}' = G''$. Otherwise, let $\deg^-(v) = q$, $\deg^+(v) = r$, and $|N_{\hat{G}'}(v) \cap A_k| = t > 0$.

$p_k(v) \geq \frac{q+t}{q+r+t}$. Removing the t vertices in $N_{\hat{G}'}(v) \cap A_k$ decreases $p_k(v)$ to $\frac{q}{q+r}$.

Now let G'' be derived from \hat{G}' by removing all edges except for those that are between A_{k-1} and A_k . (In particular, this means that for all $k \in [4]$, for all $v \in A_k \cap V(G'')$, $N_{G''}(v) \cap A_k = \emptyset$.) This decreases $p_k(v)$ for all vertices v with neighbors w such that $d(x_0, v) = d(x_0, w)$ and does not change $p_k(v)$ for all vertices v with no such neighbors. Since

$$p_{G''} = \frac{1}{|N_{G''}(x_1)|} \sum_{u \in N_{G''}(x_1)} \frac{1}{|N_{G''}(u)|} \sum_{v \in N_{G''}(u)} \frac{1}{|N_{G''}(v)|} \sum_{w \in N_{G''}(v)} p_4(w)$$

we have that $p_{G''} \leq p_{\hat{G}'}$. □

In view of Claims 1 and 2 above, we may assume that G has the following properties: $N_G(x_0) = x_1$, the only edges in G are between A_{k-1} and A_k for $k \in [4]$, and $A_k = \emptyset$ for all $k > 4$.

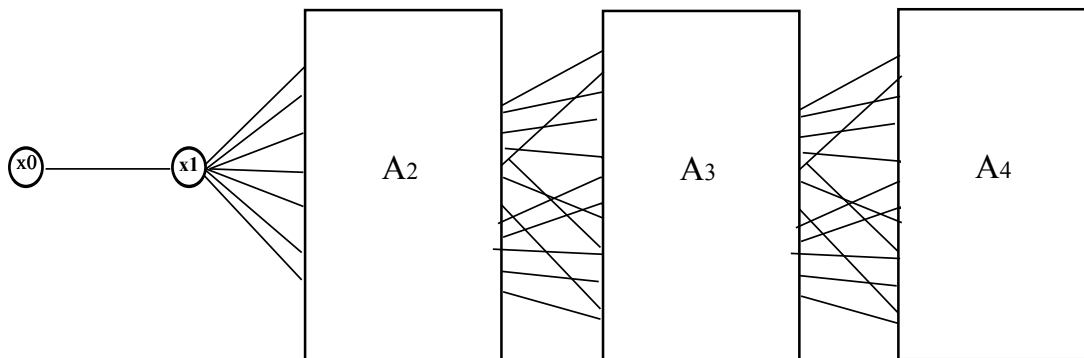


Figure 3: Our (alleged) counterexample G

Now define $G_k \subseteq G$ to be the induced subgraph of G on the vertices $A_k \cup A_{k+1}$ and let $e_k = |E(G_k)|$.

Claim 3. $e_2 > s(s-1)$.

Proof. Now we claim that the average degree of vertices in A_2 is greater than s : Let $\{d_i\}_1^{a_2}$ be the degrees of the vertices in A_2 and let $d = \sum_{i=1}^{a_2} d_i$. The vertex x_2 is chosen uniformly at random in A_2 , and the probability of stalling at a vertex with degree d_i is $\frac{1}{d_i}$. Therefore the probability of stalling at A_2 is $\frac{1}{a_2} \sum_1^{a_2} \frac{1}{d_i}$. We have

$$1/s > \frac{1}{a_2} \sum_1^{a_2} \frac{1}{d_i} = \frac{1}{H(\{d_i\})} \geq a_2/d$$

(where $H(\{d_i\})$ is the harmonic mean of the d_i). Consequently, $d/a_2 > s$. Thus the average out-degree from A_2 is greater than $s - 1$, which implies that there are more than $s(s - 1)$ edges between A_2 and A_3 . \square

Claim 4. *Let B be the subset of A_2 consisting of vertices with more than half of their outedges going to C , the subset of A_3 consisting of vertices with in-degree less than $n^{1/3}$. Let $b = |B|$ and $c = |C|$. Then $b < \frac{1}{2}a_2$.*

Proof. Define e_B to be the number of edges with one endpoint in B and the other in A_3 . Note that $c \leq a_3 < n - a_2 \leq n - 4n^{2/3}$ and consequently the number of edges with one endpoint in A_2 and the other in C is less than $n^{4/3} - 4n$. Since more than half of the outedges of each vertex in B terminate in a vertex in C , this says that $e_B < 2n^{4/3} - 8n$.

Now assume, for sake of contradiction, that $b \geq \frac{1}{2}a_2$. Then $\mathbb{P}(x_2 \in B) > 1/2$ so we have

$$\begin{aligned} 1/s &> p_G \\ &= \mathbb{P}(d(x_0, x_4) < 4 | x_2 \in B) \mathbb{P}(x_2 \in B) + \mathbb{P}(d(x_0, x_4) < 4 | x_2 \notin B) \mathbb{P}(x_2 \notin B) \\ &> \frac{1}{2} \mathbb{P}(d(x_0, x_4) < 4 | x_2 \in B) \end{aligned}$$

which says that $\mathbb{P}(d(x_0, x_4) < 4 | x_2 \in B) < \frac{2}{s}$.

Let $f = \sum_{i=1}^b d_i$ where $\{d_i\}_{i=1}^b$ are the degrees of the vertices in B . For each vertex in B , $\mathbb{P}(d(x_0, x_4) < 4 | x_2 \in B; \deg(x_2) = d_i) = \frac{1}{d_i}$. Since x_2 is chosen uniformly at random, we have

$$\begin{aligned} \frac{2}{s} > \mathbb{P}(d(x_0, x_4) < 4 | x_2 \in B) &= \frac{1}{b} \sum_{i=1}^b \frac{1}{d_i} \\ &= \frac{1}{H(\{d_i\}_1^b)} \\ &\geq \frac{1}{(1/b)f} = \frac{b}{f} \end{aligned}$$

The average out-degree from B is $\frac{f}{b} - 1$ and so we get $e_B \geq b \left(\frac{f}{b} - 1 \right) \geq \frac{s}{2} \left(\frac{s}{2} - 1 \right) = 4n^{4/3} - 2n^{2/3}$. This is a contradiction since $2n^{4/3} - 8n < 4n^{4/3} - 2n^{2/3}$.

Consequently, $b < \frac{1}{2}a_2$. □

Claim 5. *The probability that $x_3 \in A_3 \setminus C$ (given $x_3 \in A_3$) is greater than $1/4$.*

Proof. If $x_2 \in A_2$ then with probability greater than $1/2$, $x_2 \in A_2 \setminus B$. By definition, more than half of the out-edges of a vertex in $A_2 \setminus B$ terminate in $A_3 \setminus C$, and x_3 is chosen uniformly at random from the neighbors of x_2 . This yields

$$\mathbb{P}(x_3 \notin C | x_3 \in A_3) = \mathbb{P}(x_3 \notin C | x_2 \in B) \mathbb{P}(x_2 \in B) + \mathbb{P}(x_3 \notin C | x_2 \notin B) \mathbb{P}(x_2 \notin B)$$

and therefore

$$\mathbb{P}(x_3 \notin C | x_3 \in A_3) \geq \mathbb{P}(x_2 \notin B | x_2 \in A_2) \mathbb{P}(x_3 \notin C | x_2 \notin B) > (1/2)(1/2) = 1/4,$$

as desired. □

Note that $\mathbb{P}(d(x_0, x_4) < 4 | x_3 \in A_3 \setminus C) = \frac{\deg^+(x_3)}{\deg(x_3)} > \frac{n^{1/3}}{n}$. Therefore the probability of stalling at step 3 is greater than $(1/4) \frac{n^{1/3}}{n} = 1/s$, yielding a contradiction. □

Let $j = 4i + 1$ be such that the game is in Stage 3 at time j , and let x_j, y_j be the positions of the drunk and cop, respectively, after both have moved (so that it is the cop's turn). By Lemma 3.4 we have that $d(x_j, x_{j-4}) < 4$ with probability at least $\frac{1}{4}n^{-2/3}$. Consequently, since the cop had x_{j-4} as her target, we now have $d(y_j, x_j) < d(y_{j-4}, x_{j-4})$ (so the distance has decreased by at least 1) with probability at least $\frac{1}{4}n^{-2/3}$. Let Y_i be a random variable which equals the decrease in distance between time $4(i-1)$ and $4i$. Y_i is 0 with probability less than $1 - 1/s$ and is ≥ 1 with probability at least $1/s$.

Consider the 0-1 random variable X_i with $\mathbb{P}(X_i = 1) = 1/s$ for all i (note $\mathbb{E}[X_i] \leq \mathbb{E}[Y_i]$ for all i). Let $S_n = X_1 + \dots + X_n$, for all $n \in \mathbb{N}$. Consider the random process $\{X_i : i \in \mathbb{N}\}$ with the stopping rule that says the process terminates at time τ if $S_\tau = D_2 - 3$. By Wald's identity [14], $\mathbb{E}[S_\tau] = \mathbb{E}[\tau] \mathbb{E}[X_i]$. Since $\mathbb{E}[S_\tau] = \mathbb{E}[D_2] - 3$, we have that the expected stopping time $\mathbb{E}[\tau] = \frac{\mathbb{E}[D_2] - 3}{1/s}$. This is the expected number of retargetings needed to get $S_\tau = D_2 - 3$, so we have

$$\mathbb{E}[T_3] \leq 4\mathbb{E}[\tau] = 4s(\mathbb{E}[D_2] - 3) < 4((5 \log n)^{3/4} n^{1/4} - 3)(4n^{2/3})$$

Stage 3 terminates when the distance between the cop and the drunk is less than four, and it is the cop's turn. In Stage 4, which terminates when the drunk is captured, the cop uses the greedy strategy, defined as follows. Suppose that the strategy enters Stage 4 at time t , during which time the drunk is at vertex x_t and the cop is about to move from vertex y_{t-1} . Then $d(x_t, y_{t-1}) \leq 3$, and the cop moves such that $d(x_t, y_t) \leq 2$. Now for any $r > t$, if the drunk moves such that $d(x_r, y_{r-1}) = 3$, the cop can choose y_r to ensure

that $d(x_r, y_r) = 2$. For each r , with probability at least $1/\Delta$, the drunk moves “toward” the cop—i.e., such that $d(x_r, y_{r-1}) = 1$; if that happens, the cop can choose $y_r = x_r$, capturing the drunk. This takes at most Δ expected moves, so $\mathbb{E}[T_4] \leq \Delta$ where T_4 is the expected time spent in Stage 4.

Adding together our results about the expected time to complete each of the four stages yields the following bound on the expected capture time:

$$\begin{aligned} \sum_{i=1}^4 \mathbb{E}[T_i] &\leq \text{diam}(G) + \mathbb{E}[D_1] + 16n^{2/3}(\mathbb{E}[D_2] - 3) + \Delta \\ &< \text{diam}(G) + 1 + \sqrt{n(1 + 5 \log n)} + 4((5 \log n)^{3/4}n^{1/4} - 3)(4n^{2/3}) + \Delta \\ &= \text{diam}(G) + \Delta + o(n) \end{aligned}$$

In fact, we can bound $\text{diam}(G) + \Delta$ with a bit of graph theory.

Lemma 3.5. *For any graph G with $|V(G)| = n$, $\text{diam}(G) + \Delta \leq n + 1$.*

Proof. Assume, for sake of contradiction, that there is a graph G such that $\text{diam}(G) > n - \Delta + 1$. Let $u, v, w \in V(G)$ be (not necessarily distinct) vertices in G such that $\deg(v) = \Delta$ and $d(u, w) = d \geq n - \Delta + 2$. Now we break this proof into two cases:

Case 1: v lies on a shortest path between u and w .

Let P_1 be a shortest $u - w$ path containing v . At most two neighbors of v may lie on P_1 , so there are at least $\Delta - 2$ vertices not on P_1 . Since the length of P_1 is at least $\geq n - \Delta + 2$, there are at least $n - \Delta + 3$ vertices in P_1 . But now we have that $|V(G)| \geq \Delta - 2 + n - \Delta + 3 > n$, which is a contradiction.

Case 2: v is not on any shortest $u - w$ path.

Let P_2 be a shortest $u - w$ path. If more than 2 neighbors of v are in P_2 , then v is also on a shortest $u - w$ path (let x_1, x_2 , and x_3 be the neighbors of v on P_2 , appearing in that order; then the section involving the three neighbors of v could be replaced with $x_1 - v - x_3$ to create another shortest $u - w$ path). Therefore there are at least $\Delta - 1$ vertices not on P_2 (v and $\Delta - 2$ of its neighbors), and at least $n - \Delta + 3$ vertices on this path. So once again, $|V(G)| \geq \Delta - 1 + n - \Delta + 3 > n$, which is a contradiction.

Therefore $\text{diam}(G) \leq n - \Delta + 1$ for all graphs G . □

Therefore we have the following theorem about the expected capture time.

Theorem 3.6. *On a connected, undirected, simple graph on n vertices, a cop with the described four-stage strategy will capture a drunk in expected time $n + o(n)$.*

4 Open Problems

The reader may, for instance, have noticed that in the “ladder to the basement” example of Section 2, we considered a cop who was not only greedy but also rather insistently foolish. What about the greedy cop who makes distance-minimizing decisions at random? The “ladder to the basement” graph is no longer a problem for her, (the expected capture time

in this example is now less than n). Is it possible that the greedy algorithm with disputes settled by a random decision between choices is enough to guarantee time $n + o(n)$?

It is also possible that a deterministic greedy cop who breaks ties by considering her distance to vertices previously occupied by the drunk will capture in expected time at most $n + o(n)$.

An alternative greedy strategy, suggested by Andrew Beveridge [1], concerns itself with minimizing the drunk's expected hitting time to the cop at every step. It would be interesting to see if this strategy also has expected capture time at most $n + o(n)$.

Acknowledgments

This work has benefited from conversations at Microsoft Research, in Redmond, Washington, with Omer Angel, Ander Holroyd, Russ Lyons, Yuval Peres, and David Wilson.

References

- [1] A. Beveridge, private communication, 2nd Annual GRASCan Workshop, April 2013.
- [2] G. R. Brightwell and P. Winkler, Maximum hitting time for random walks on graphs, *J. Random Structures and Algorithms* **1** #3 (1990), 263–276.
- [3] G.R. Brightwell and P. Winkler, Gibbs measures and dismantlable graphs, *J. Comb. Theory (Series B)* **78** (2000), 141–169.
- [4] D. Coppersmith, P. Tetali and P. Winkler, Collisions among random walks on a graph, *SIAM J. Disc. Math.* **6** #3 (1993), 363–374.
- [5] M. Hall, Projective planes, *Trans. Amer. Math. Soc.* **54** (1943), 229–277.
- [6] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* **30** (1906), 175–193.
- [7] G. MacGillivray, private communication, ca. May 2011.
- [8] R. J. Nowakowski and P. Winkler, Vertex to vertex pursuit in a graph, *Discrete Math.* **43** (1983), 235–239.
- [9] R. Peyre, A probabilistic approach to Carne's bound, *Potential Anal.* **29** (2008) #1, 17–36.
- [10] A. Quilliot, Thesis, Université de Paris VI, 1983.
- [11] J. Soares, Maximum diameter of regular digraphs, *J. Graph Theory* **16** (1992), 437–450.
- [12] N. Th. Varopoulos, Long range estimations for Markov chains, *Bull. Sci. Math.* **109** (1985) 225–252.
- [13] O. Veblen and W. H. Bussey, Finite projective geometries, *Trans. Amer. Math. Soc.* **7** (1906), 241–259.
- [14] A. Wald, On cumulative sums of random variables, *The Annals of Mathematical Statistics* **15** #3 (1944) 283–296.

- [15] A. Berarducci and B. Intrigila, On the cop number of a graph. *Adv. Appl. Math.* **14** (1993) 389–403.
- [16] A. Bonato, P. Golovach, G. Hahn, and J. Kratochvil, The capture time of a graph, *Discrete Math.* **309** (2009) 5588–5595.
- [17] G. Hahn, F. Laviolette, N. Sauer and R.E. Woodrow, On cop-win graphs, *Discrete Math.* **258** (2002) 27–41.
- [18] G. Hahn and G. MacGillivray, A note on k-cop, l robber games on graphs, *Discrete Mathematics* **306** (2006) 2492–2497.