,
Criminal Pursuit on a Preferential Attachment Tree

Charlie Z. Marshak
UCLA Department of Mathematics

M. Puck Rombach, Andrea L. Bertozzi
UCLA Department of Mathematics
and Institute of Pure and Applied Mathematics

Maria R. D'Orsogna
UCLA Department of Mathematics
and CSUN Department of Mathematics
(MURI Collobaration)
(Dated: August 12, 2014)

We model the hierarchal evolution of an organized crime network under the competing influences of a kingpin and the police. At each time step, the network's members incorporate new criminals into the network. Additionally, an officer looks to arrest as many criminals, old and new, as she can, with the kingpin's arrest being the ultimate goal. The growth process follows a preferential attachment model: criminals arrive at a constant rate and attach to existing criminals with probability roughly proportional to the inverse distance to the lowest level of the hierarchy. Here, the lowest ranking criminals are those without any underlings. Then, the officer follows a self-avoiding random walk and she can arrest criminals at any given moment in her pursuit. Alternating between growth and pursuit fashions our dynamic graph model of an organized crime network. For the growth process, we study the evolution of criminal rankings in the network and the degree distribution in the network as $t \to \infty$. For the pursuit process, we numerically simulate the dynamic game and infer that a degree-threshold strategy for the police is optimal given that the officer's knowledge is confined to local network data.

INTRODUCTION

Network science has become a central tool in the modeling of criminal behavior [1–3]. Haeurt and his collaborators paired networks and evolutionary game theory to model human cooperation within a Darwinian framework [4–6]. The joint work of the last author with Short and Brantingham employed similar ideas to model the interplay of criminals and informants in society [7, 8], and then with the addition of a social network apparati by Short, et al. [9]. Using these models as a springboard, we model the dynamics of a hierarchal crime network that undergoes expansion via recruitment and then shrinkage due to arrest. The tree structure is an appropriate choice for networks that are large and entrepreneurial in nature [10, 11]. Our model has two key components: network growth and police pursuit.

The first component of the process, the growth, introduces a preferential attachment mechanism that follows in spirit of the mathematical work of Szymański [12], Yule [13], Barabási-Albert [14], and Mahmoud [15, 16]. This growth process models criminal recruitment. Criminals are added to the network at a constant rate and each form a unique connection to the existing network. In criminology, recruitment into a criminal network is a complicated process that involves both the recruiter and the network to acquire trust and to engage in frequent communication [17, 18]. In this mathematical simulation of recruitment, we model the recruitment probabilistically and attempt to capture some of the human elements. We assume the incoming criminals will most likely form a link to those with closest ties to street activity because criminological studies suggest prospective members of a criminal network are those already engaged in illicit street activities themselves [17]. Also, those criminals without underlings will have more more time to communicate to prospective criminals and are likely the most visible to people outside of the network. Again, proximity to street crime in our network model is understood in terms of the minimum distance a criminal is from any of the other criminals without underlings.

We remark, that in our model, while street criminals are the most likely to recruit, even the kingpin can form a link to a prospective throughout the life of the network, though the probability this occurs will become small as the network grows large. Due to the possibility of investigation and arrest, which will be formally introduced as a graph game later, criminals benefit from maintaining a buffer between themselves and street criminals in the network. The more recruiting any criminal does, the more vulnerable and visible she is to police capture. Thus, long-time members of the network should have less incentive to recruit, which is reflected in their low probability of adding a link.

The second component of the process, the pursuit, adapts the classic "birth-death" apparatus from population dynamics [19, 20] into this criminal network model. We detail the movement of a single officer whose ultimate goal is to capture the kingpin. The officer proceeds by a self-avoiding random walk beginning with one of criminals without underlings. If she arrives at the kingpin, then the pursuit is successful and the process is terminated. If the officer can no longer move, then the pursuit is unsuccessful and the officer must attempt to capture the kingpin the next round. At any point in the investigation, the officer can arrest the criminal she is currently investigating and all of the criminals below. By removing criminals from the network, the officer makes the kingpin more vulnerable to future capture.

We view the process as a game on a graph with the two players represented by a kingpin and an officer. The kingpin wins if the officer does not capture the kingpin in finite time and the officer wins otherwise. We study strategies for arrest that the officer can pursue.

## PART I: PREFERENTIAL ATTACHMENT

Let $T^{(0)}$ be the initial criminal network with unique kingpin. Set $T^{(t)}$ to be a recursively defined network at time t with $V^{(t)}$ the set of criminals. The set of criminals $L^{(t)}$ contained in $V^{(t)}$ denotes those criminals without any underlings. We assume $L^{(t)}$ are the criminals closest to street activity. To obtain $T^{(t)}$ from $T^{(t-1)}$, we add k new criminals to $T^{(t-1)}$; we call k the rate of new criminal arrival . Let $\text{dist}_t(j)$ denote the minimum distance from criminal j to $L^{(t)}$. Each new criminal forms a link to an existing criminal j on $T^{(t-1)}$ with probability proportional to weight $w_{t-1}(j)$. In particular, if a new criminal i is added to the tree $T^{(t)}$, then the probability that an edge e_{ij} from i to j is formed is given by:

$$\mathbb{P}_t(e_{ij}) = \frac{w_{t-1}(j)}{\sum_{j'} w_{t-1}(j')} \tag{1}$$

The weight $w_t(j)$ of criminal j at time t is defined as:

$$w_t(j) := \frac{1}{\text{dist}_t(j) + A} \tag{2}$$

with $A \in \mathbf{R}$. We set $A = 1$ for this model. Observe that $0 \le w_t(j) \le 1$ for all $j \in V^{(t)}$ and the maximum of w_t is attained for $j \in L^{(t)}$. See Figure [1] for an illustration of the probability assigned to each criminal for a given experiment. We remark that as $A \to \infty$, $w_t(j)$ is a uniform weight on each of the nodes and the process becomes the recursive trees found in [12, 15, 16]. As $A \to 0$, criminals with higher and higher probability link to criminals without underlings and the model will become increasingly deterministic. Our decision to use $A = 1$ stems from the fact that we wish to have a middle ground between these two extremes: more weight should be given to criminals
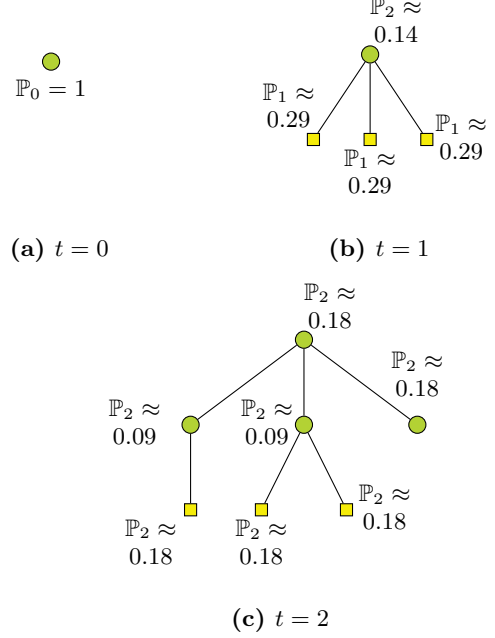
Figure 1: The above figure shows $\mathbb{P}_{t+1}$, the probability a new link is formed to one of the k new criminals entering the network . Here, the rate of criminal arrival k is 3 and $T^{(0)}$ is the kingpin alone. To calculate $\mathbb{P}_{t+1}$, we must find d_t, the distance a criminal is from $L^{(t)}$ and then w_t, the weight of the criminal i at t given by $\frac{1}{d_t+1}$. Lighter green leaves denotes criminals new to the network on the given time step.

without underlings, while still permitting long-standing criminals to have a chance of forming a new criminal link.

Degree Distribution

The Barabási-Albert model is a preferential attachment model, where nodes are added with greatest probability to network nodes of high degree [14, 21] and is summarized as a network where "the rich get richer". A salient mathematical feature of this preferential attachment model is its heavy-tailed degree distribution, which obeys a power-law [21]. From the data seen in Figure [2], we conjecture the degrees in this model are also heavy-tailed, and further, obey an exponential law with a shift. While the empirical exponential law is difficult to numerically capture via numerical experiments [22], we argue why there should be such a law heuristically. When criminal adds a new link, then effectively a new network process begins. This self-similarity permits a high probability that criminal degrees remain at a fixed proportion throughout the life of the network.

This Model's Terminology	Network Theory Terminology
The kingpin	The root
$T^{(t)}$: Criminal Network	A rooted tree at time t
$V^{(t)}$: Criminals	Verticies or nodes at time t
$L^{(t)}$: Criminals without underlings or street criminals	The set of leaves in the network at time t

Table I: Above, we provide a dictionary of the criminal network terminology used in this paper and that of standard network theory.

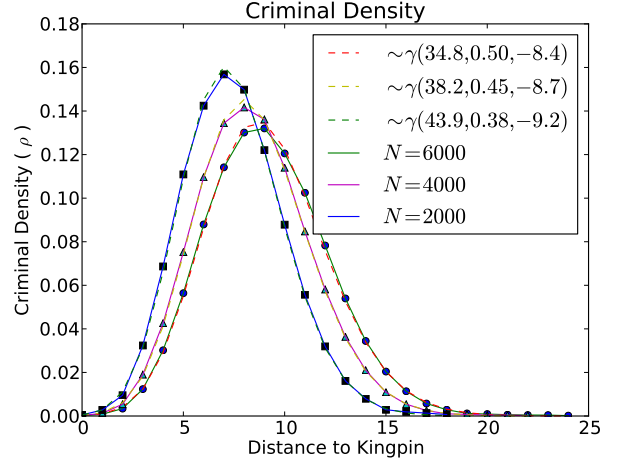| Parameter	Description
t	Time
k	New criminals added to the network at each timestep
$T^{(0)}$	Initial network configuration
A	Parameter in $w_t(j)$. We set $A = 1$.
N	Number of criminals before growth experiment stopped

Table II: We record the parameters of the growth model above.
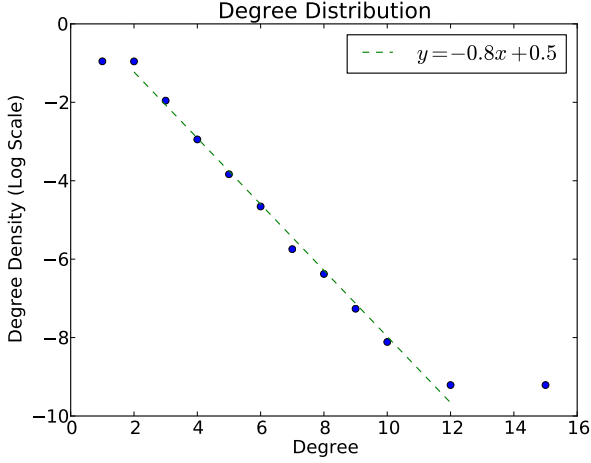
Figure 2: The Criminal Network was grown to 10000 (N). 500 experiments were performed, averaged, and then collected into a histogram. Observe that the number criminal with 1 and 2 degree are roughly at the same proportion. The criminals of degree 1 are precisely the criminals without underlings, or street criminals. The criminals of degree 2 are with high probability those criminals a distance 1 from the aforementioned street criminals. These two populations represent the pipeline into the network and it makes sense that these two populations occupy the same proportion for all time. The tail of the distribution, those criminals with high degree, picks up noise on account of its rarity. The fit becomes more and more pronounced as $N \to \infty$ and the number of experiments increases.

Figure 3: The number of criminals N in the network simulations were 2000, 4000; and 6000. We fixed the arrival rate to be $k = 4$. Our data is generated from 20 experiments. $T^{(0)}$ was a complete tertiary tree of height 3. We fit the data with a shifted γ-density.

Criminal Density

We found the criminal density depends on their distance to the kingpin. Figure [3] suggested a shifted gamma (γ-) density numerically approximates this relationship. As $N$ grows, the expected distance from the kingpin increases, which align's with the kingpin's desire to ensure a buffer between the possible police investigations of street criminals when the graph game begins. The probability density of a γ-distributed random variable is given by:

$$\rho_{\alpha,\beta,s}(h) = \frac{\beta^\alpha}{\Gamma(\alpha)}(h-s)^{\alpha-1}e^{-\beta(h-s)} \quad h \geq 0. \quad (3)$$

and if X = distance to kingpin of a randomly selected criminal, then

$$\mathbb{P}(X \leq h_0) \approx \int_0^{h_0} \rho_{\alpha,\beta,s}(h) \, dh \quad (4)$$

Above, h and h_0 denote distances to the kingpin.
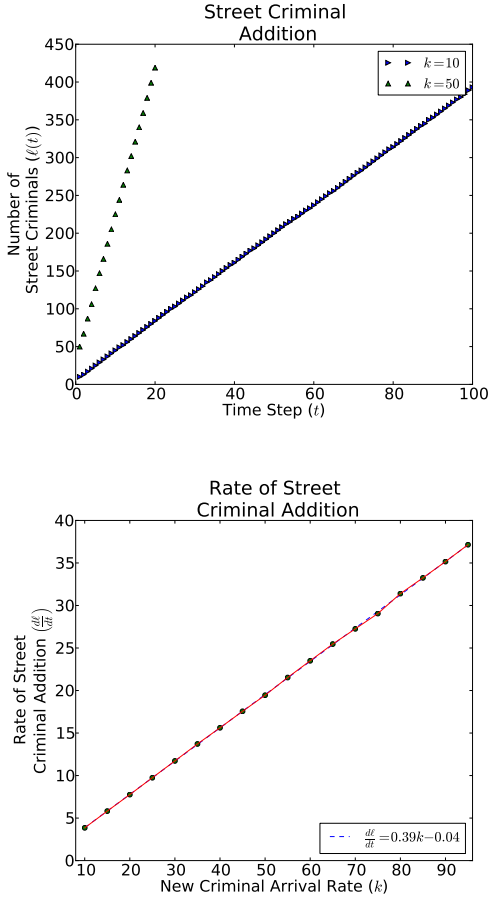
Figure 4: The number of criminals was fixed to be $N = 1000$, while k was varied. The function $\ell(t)$ appears linear; moreover, the slope of this relation appears to linearly depend on k, i.e. $d\ell/dt = c_k \approx .39k$ (bottom).

The Growth of Street Criminals

Street criminals are those criminals without any underlings. They have the highest probability of recruiting new criminals. We assume street criminals are those primarily responsible for personally carrying out the illegal activity that permit the network to profit and whose presence make the streets more dangerous to network outsiders. From a modeling perspective, we expect as the network grows, so too should the aggregate number of street criminals. Set $\ell(t) = |L^{(t)}| =$ the number of street criminals at time t. Then, Figure 4 suggests that $\ell(t)$ is linear in t and that $\frac{d\ell}{dt} = c_k$ is a constant that depends on k alone. The positivity agrees with our expectation that street criminals grow as the network does. It also agrees with the fact that criminals of degree 1 remain at fixed proportion through the life of the network as the rate of criminal arrival k is constant (see Figure 2).
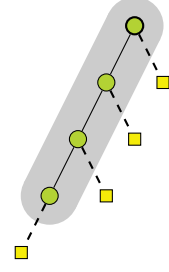
Figure 5: Above the topmost criminal has a path that realizes D_t, the maximum distance to $L^{(t)}$; this path is highlighted in grey. The light dashed edges denote possible connections a new criminal can make on this path. Suppose the length of the path is n. The probability the path will increase by 1 is $\frac{1}{1 + \frac{1}{2} + \ldots + \frac{1}{n}}$.

Network Proximity to Street Criminals

The model has the feature that each criminal in the network is never too far removed from street activity. In other words, all criminals have at least one path that connects them to another criminal in $L^{(t)}$. This effectively gives even the longest standing criminals in the network some access to the street as well as makes all criminals vulnerable to police intervention. We can view

$$D_t := \max_{j \in V^{(t)}} d_t(j) \tag{5}$$

as a random variable on $T^{(t)}$ for any t.

We predict D_t to be bounded in the limit; that is, there exists an $M > 0$ such that

$$\lim_{t \to \infty} \mathbb{P}(D_t \leq M) = 1 \tag{6}$$

assuming the limit exists. As indicated by Figure 5,

$$P(D_t > n) \leq \frac{1}{1 + \ldots + \frac{1}{n}} \tag{7}$$

for a particular $T^{(t)}$. The boundedness is further inspected by numerical experiments found in Figure 6.

PART II: ADDING POLICE PURSUIT

Once police pursuit is incorporated into the dynamics above, we view our process as a game with two players: the network led by a kingpin and the police led by a single officer. A win for the officer is when she catches the kingpin in finite time. A win for the network is precisely the opposite.

The Model with Police

Let $T^{(0)}$ be the criminal network at $t = 0$, one sufficiently large so that a win for the police is not inevitable.
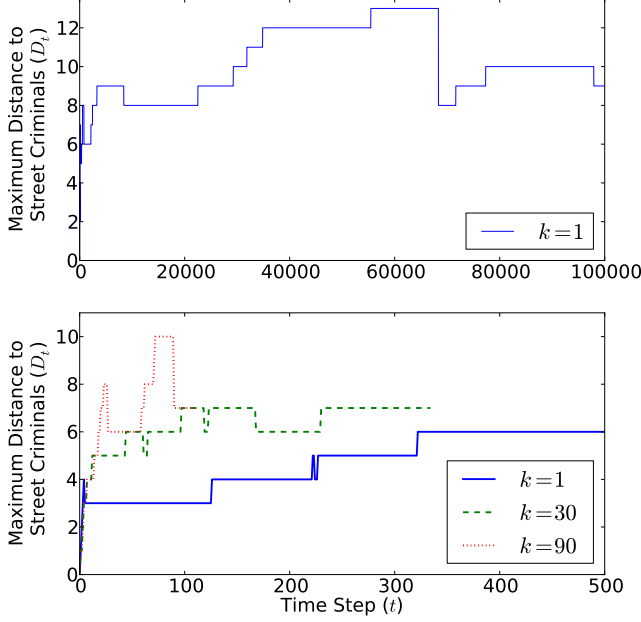
Figure 6: Above, we compute the maximum distance any criminal in the network is to a street criminal. We inspect various criminal arrival rates k. We see that for larger k, there is a sharper increase initial increase, though the maximum distance remains bounded for criminal networks of 10^5 (left) and 10^4 (right).

We define a recursive process to obtain $T^{(t)}$ from $T^{(t-1)}$ in two sequential steps.

Step 1 : This is the police's move. We place the officer uniformly at random on one of the street criminals. The officer can either investigate or arrest. The officer investigates by moving randomly among criminals linked to the current location of the officer such that the officer never travels to the same criminal twice. The greater the degree of the current criminal the greater the variability of where the officer could end up. Arresting consists of removing the criminal that the officer is currently investigating and all of the criminals below. The game ends if the officer reaches the kingpin. The pursuit ends and the game continues if the officer reaches another leaf or chooses to arrest the current criminal.

Step 2 : This is the network's move. We conduct a round of preferential attachment with k new criminals described in the previous section.

The officer has a choice to make at each time step in order to optimize her chances of winning: when should she arrest if at all? Strategies are elaborated in what follows.

Strategies and the Beat(S)

During the police's move, each new investigation brings a choice: to investigate another criminal or to arrest the current one. We assume the officer does not know the global structure of the network save that the network is hierarchal in structure. As she investigates the network, her information is limited to local data and the information accrued during the move, meaning she is able to view a criminal's connections (i.e. the degree) and record all of the criminals she has previously investigated. After the round is over, the officer looses the information she has amassed and must start anew the next round.

Here are the three strategies the officer can pursue assuming that her knowledge is limited to the above:

(1) The officer investigates p times and then arrest; this strategy will be denoted by $S_A(p)$.

(2) The officer always investigates. She is successful only if she reaches the root; this strategy will be denoted by S_I.

(3) The officer investigates unless reaching a criminal of degree at least q, in which case she arrests; this strategy will be denoted by $S_D(q)$.

Figure 7 shows $S_A(1)$, S_I, and $S_D(4)$ after the initial criminal configuration.

Let S be the officer's strategy determining when she arrests a criminal during her investigation of the network. We assess a strategy by calculating the maximum criminal arrival rate such that the officer wins with probability 1. Let:

$$\text{Beat}(S) := \max \left(k \left| \begin{matrix} \text{police eliminate network in} \\ \text{finite time with probability 1} \end{matrix} \right. \right) \qquad (8)$$

The greater Beat(S), the more effective a strategy S is against a growing criminal network. See Figure [8] for examples.

I. OPTIMAL STRATEGY

Roughly speaking,

$$S_I = \lim_{p \to \infty} S_A(p) = \lim_{q \to \infty} S_D(q) \qquad (9)$$

For strategy $S_A(p)$, this says that as the number of predetermined investigations grows this grows in similarity to the investigation only strategy S_I. Similarly for $S_D(q)$, as the degree threshold increases, the officer pursues an investigation only strategy. Depending on the initial condition q for $S_D(q)$ can be determined.
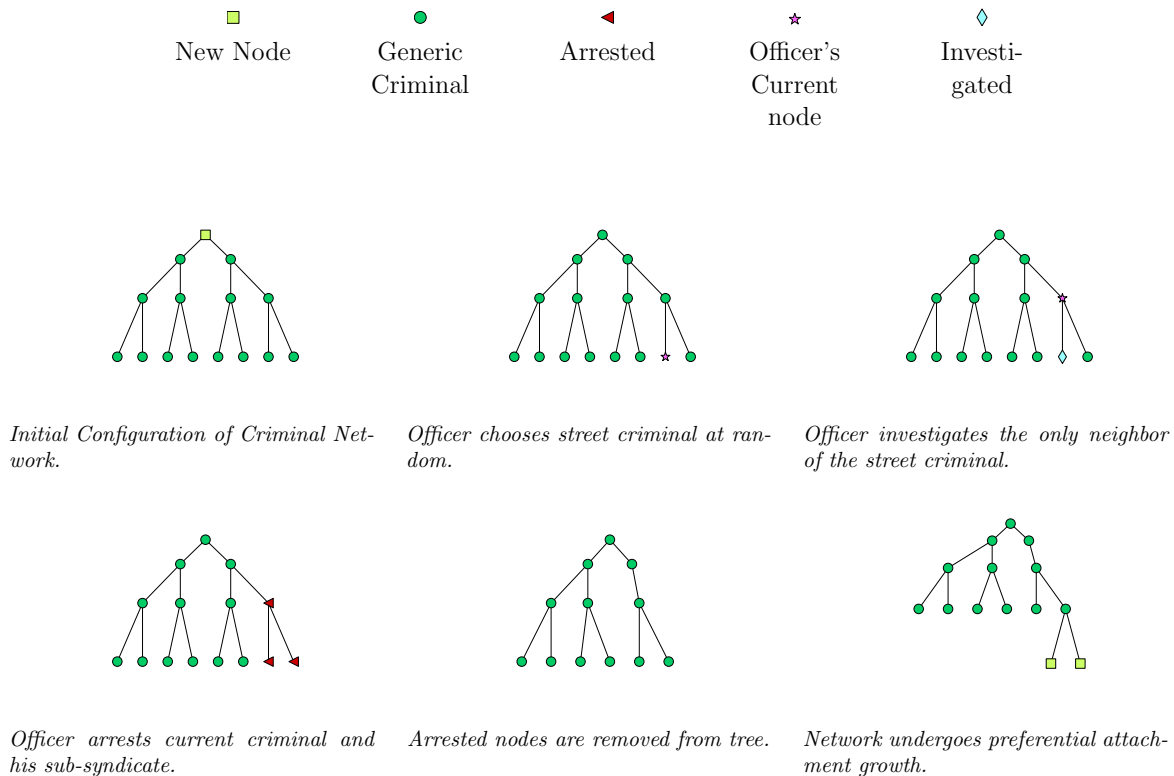
Figure 7: Above shows one iteration of the coupled pursuit and growth where the officer pursues strategy $S_A(1)$. This strategy is when the officer investigates once and then arrests. Here $k = 2$ for the growth process.

II.   CODE

We used NetworkX, NumPy, and SciPy for all the simulations. Matplotlib and tike were used for the graphical representations. The code for the preferential attachment tree, the dynamic game, and the experiments can be found at `https://github.com/cmarshak/GameOfPablos`.

ACKNOWLEDGEMENTS

[1] J. S. McIllwain, Crime, Law and Social Change **32**, 301 (1999).
[2] C. R. Browning, R. D. Dietz, and S. L. Feinberg, Social Forces **83**, 503 (2004).
[3] N. F. Johnson, C. Xu, Z. Zhao, N. Ducheneaut, N. Yee, G. Tita, and P. M. Hui, Physical Review E **79**, 066117 (2009).
[4] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund, Science **316**, 1905 (2007).
[5] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, Nature **441**, 502 (2006).
[6] C. Hauert and G. Szabó, American Journal of Physics **73**, 405 (2005).
[7] M. B. Short, P. J. Brantingham, and M. R. D'Orsogna, Phys. Rev. E **82**, 066114 (2010), URL `http://link.aps.org/doi/10.1103/PhysRevE.82.066114`.
[8] M. McBride, R. Kendall, M. B. Short, and M. R. DOrsogna, UCI Department of Economics Working Paper pp. 12–13 (2012).
[9] S. G. McCalla, M. B. Short, and P. J. Brantingham, Journal of Statistical Physics **151**, 673 (2013).
[10] M. Kenney, Survival **45**, 187 (2003).
[11] W. L. Reed and S. H. Decker (2002).
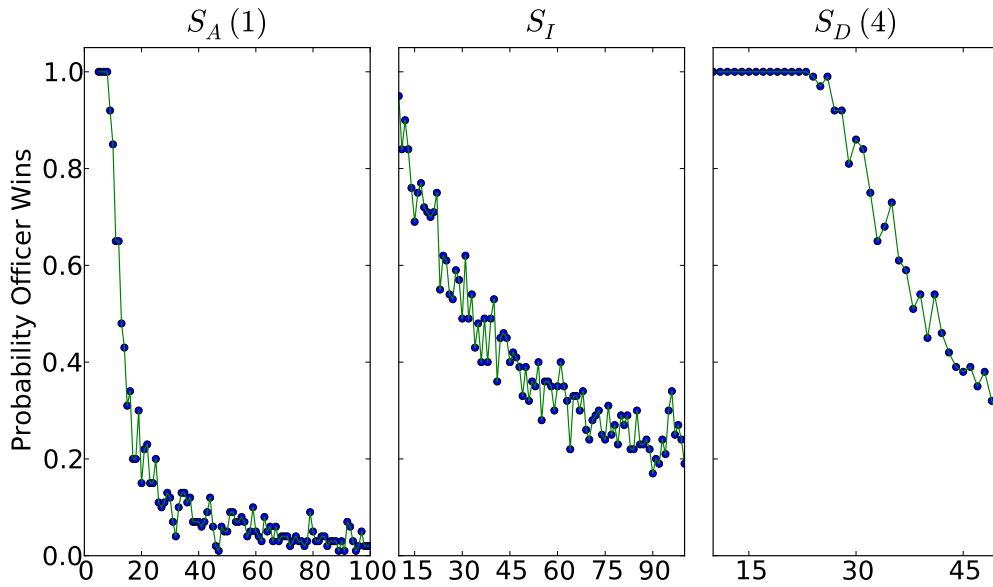
Figure 8: We numerically determined the probability of winning with each strategy using 100 trials. From left to right: $S_A(1)$, S_I, and $S_D(4)$. We infer that $S_D(4)$ has the highest Beat and it's the best strategy of those inspected. Note that S_I does not have a k for which the probability of winning is nearly 0 as $S_A(1)$ does.

[12] J. Szymański, North-Holland Mathematics Studies **144**, 297 (1987).

[13] G. U. Yule, Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character **213**, 21 (1925).

[14] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[15] R. T. Smythe and H. M. Mahmoud, Theory of Probability and Mathematical Statistics **51**, 1 (1995).

[16] H. M. Mahmoud, Journal of Computational and Applied Mathematics **41**, 237 (1992).

[17] J. A. Densley, Social Problems **59**, 301 (2012).

[18] H. H. Valeria Pizzini-Gambetta (2011), eCPR General Conference.

[19] D. G. Kendall, The Annals of Mathematical Statistics pp. 1–15 (1948).

[20] C. W. Gardiner et al., *Handbook of Stochastic Methods*, vol. 3 (Springer Berlin, 1985).

[21] B. Bollobás, O. Riordan, J. Spencer, G. Tusnády, et al., Random Structures & Algorithms **18**, 279 (2001).

[22] A. Clauset, C. R. Shalizi, and M. E. Newman, SIAM review **51**, 661 (2009).

[23] R. Durrett, *Random Graph Dynamics* (Cambridge University Press, 2007), ISBN 978-0-521-86656-9, URL http://www.cambridge.org/9780521866569.

[24] D. G. Kendall, Journal of the London Mathematical Society **1**, 385 (1966).

[25] S. Lalley, class Notes at http://galton.uchicago.edu/~lalley/Courses/312/Branching.pdf.

[26] F. Galton and H. Watson, J. Anthropol. Inst. **4**, 138 (1874).

[27] T. Luczak, Random Structures & Algorithms **1**, 287 (1990).

[28] B. Bollobás, Transactions of the American Mathematical Society **286**, 257 (1984).

[29] D. Aldous and J. Pitman, in *Annales de l'Institut Henri Poincare (B) Probability and Statistics* (Elsevier, 1998), vol. 34, pp. 637–686.

[30] B. Bollobás, K. Gunderson, C. Holmgren, S. Janson, and M. Przykucki, Electronic Journal of Probability **19**, 1 (2014).

[31] J.-P. Dion, Journal of Applied Probability pp. 687–694 (1974).

[32] D. M. Hull, Theoretical Population Biology **54**, 105 (1998).

[33] D. Aldous, Microsurveys in Discrete Probability **41**, 1 (1998).

[34] F. T. Bruss, Journal of applied probability pp. 915–919 (1984).

[35] M. Quine, Journal of Applied Probability pp. 411–422 (1970).

[36] LANL, *Networkx*, BSD Open Software Liscence, http://networkx.github.io/.

[37] C. Marshak, *Pruneup*, Git repository, https://github.com/cmarshak/PruneUp.

[38] N. Alon and J. H. Spencer, *The Probabilistic Method* (John Wiley & Sons, 2004).