NTNU
Norwegian University of Science and Technology

Q2S

Colonel Blotto in the Phishing War

Pern Hui Chia

Centre for Quantifiable Quality of Service in Comm. Systems (Q2S), NTNU

John Chuang

School of Information, UC Berkeley

GameSec 2011, Nov 14-15, College Park, Maryland, USA

Outline

- Background
  - Phishing
  - Colonel Blotto
- Modeling : Colonel Blotto Phishing game
- Analysis
- Implications to Anti-Phishing

Background

Background:

Phishing

- Annual Phishing losses?
  - $15.6 billion in identity theft loss [FTC 2006]
  - $3.2 billion in phishing loss [Gartner 2007]
  - $61 million (with ~0.2% actual victim rate, $200 median loss) [8]
- Characteristics:
  - ~30,000 attacks per 6-month [APWG]
  - Weak vs. strong phisher (e.g., Rock-Phish & Avalanche)
  - Different ways to host a phish (e.g., compromised servers, free-hosting services)
  - Can be hard to take down (e.g., Rock-Phish & Avalanche use fast-flux IP switching)
  - Not all phishes detected (information asymmetry)
- Q: What is the optimal strategy of a phisher?

NTNU Det skapende universitet

Background:

Colonel Blotto game

- 2-player constant-sum
- Allocation of finite resources in *n* battlefields
- Borel (1921)
- Borel and Ville (1938) : symmetric resources, n=3
- Gross and Wagner (1950) : asymmetric resources, but solved n=2 only

.. [complex, lack of pure strategies] ..

- Roberson (2006) : characterization of unique equilibrium payoff

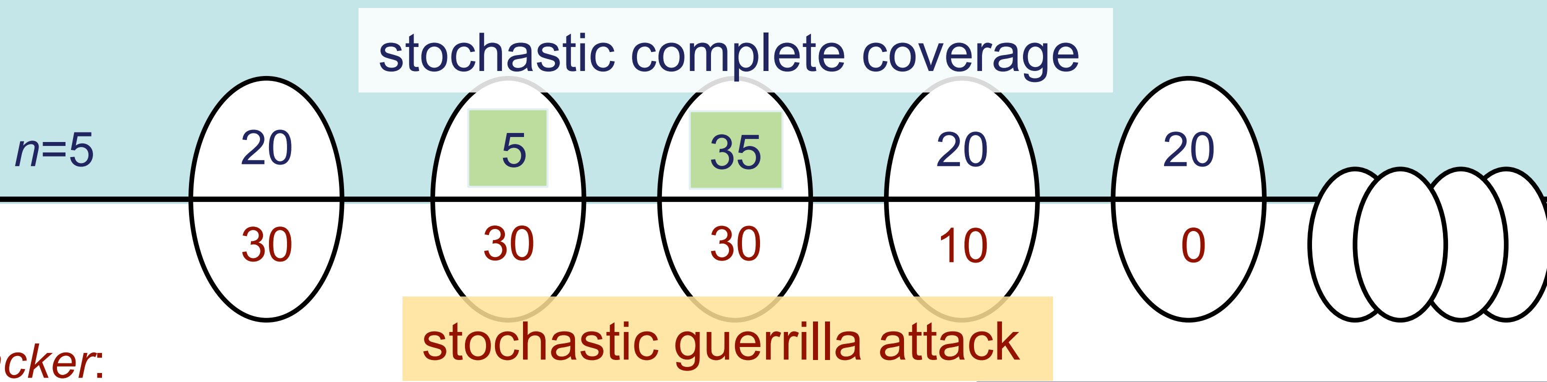
Background: Colonel Blotto game

Application to Security?

Information asymmetry?

Colonel Blotto:

Limited resource = 100 soldiers



Attacker:

Symmetrical resource = 100

Asymmetrical resource < 20 (trivial)

Asymmetrical resource > 20 (complex!)

Kovenock et al. (2010):
- endogenous dimensionality

Roberson (2006):
- payoff w.r.t. resource asymmetry

Modeling : Colonel Blotto Phishing (CBP)

Modeling:

Colonel Blotto Phishing game

- Player: takedown company vs. phisher
- Battlefield: a phish
- Objective: maximize (minimize) fraction of phishes with more than a certain uptime
- Resource: infrastructure, manpower, time (finite) (use it or lose it) (defender has more resources)
- Cost: <u>low</u>: use a free-hosting service
  <u>medium</u>: register a new domain
  <u>high</u>: compromise a server

- Stage: (1) create – detect
  (2) resist – takedown
- Can phisher win in a detected battlefield?
  - No, if phisher's resource is much lower (total lock-down)
  - Yes, if phish survives a certain uptime
    - Not resolving phish URL at every access, or temporarily removing a phish [6]
    - Re-compromising a vulnerable server [7]
    - Fast-flux IP switching (e.g., by Rock-Phish & Avalanche)

Phisher: How many new phishes to create?

S1

detected phishes

undetected phishes

cost

S2

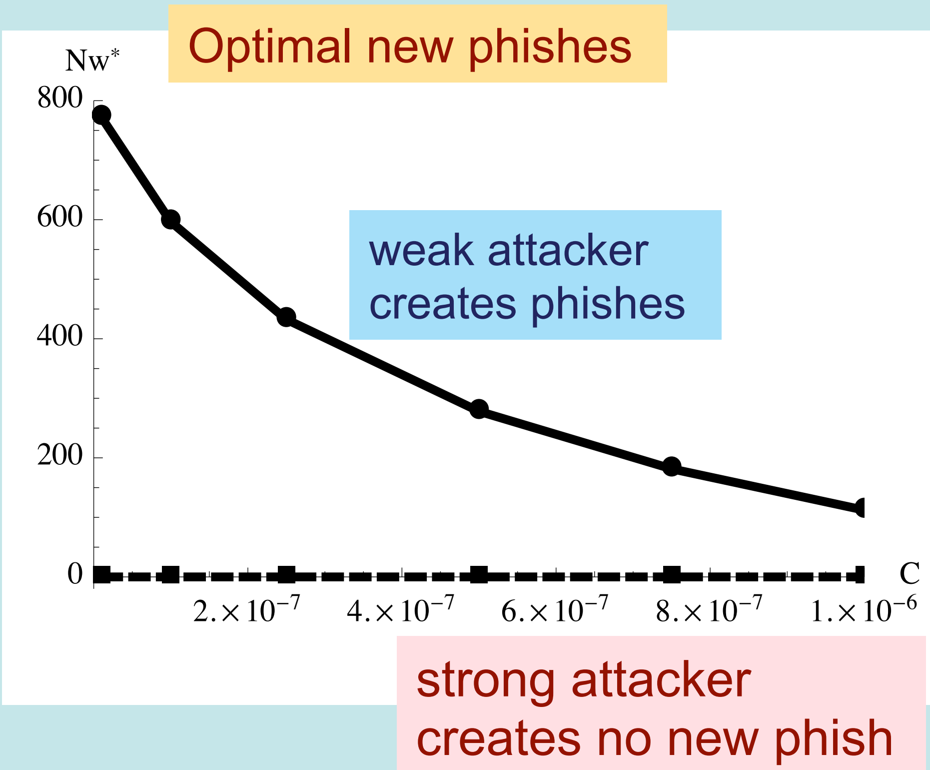
$$E(\pi_w) = \begin{cases} \dfrac{R_w}{2R_s} & \text{if} \quad 1 \geq \dfrac{R_w}{R_s} \geq \dfrac{2}{n_d} \\ \dfrac{2}{n_d} - \dfrac{2R_s}{(n_d)^2 R_w} & \text{if} \quad \dfrac{2}{n_d} \geq \dfrac{R_w}{R_s} \geq \dfrac{1}{n_d - 1} \\ 0 & \text{if} \quad \dfrac{1}{n_d} \geq \dfrac{R_w}{R_s} \end{cases}$$

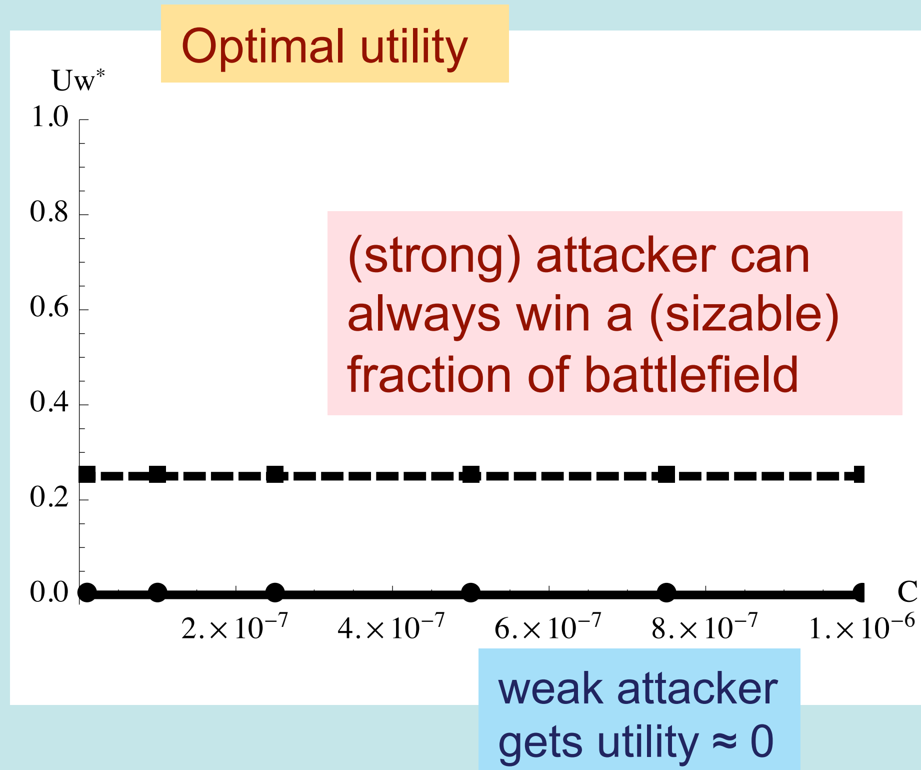
Roberson (2006)

Analysis Results

Phisher's strategy C1:

Perfect Detection (same settings as in [4])

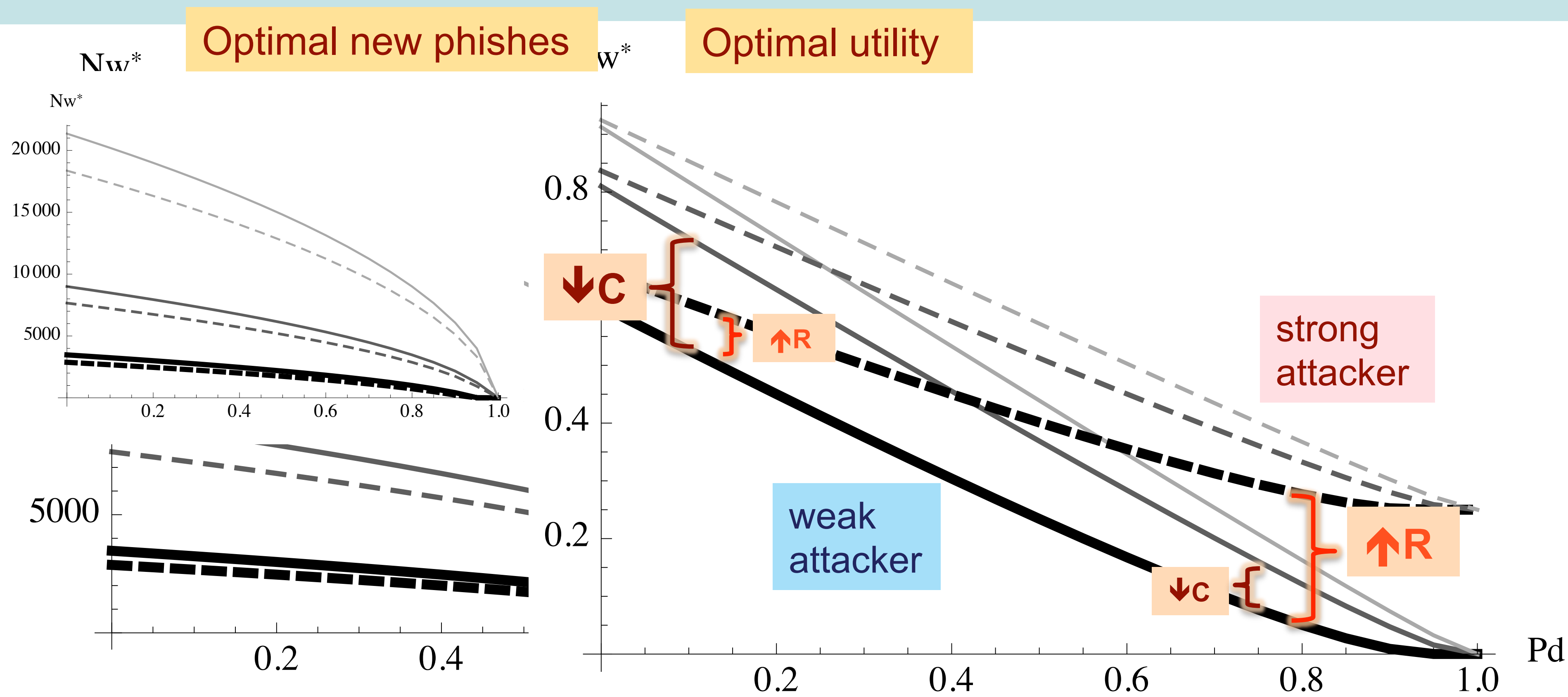




- Resource asymmetry: strong attacker vs. defender = 1/2
  weak attacker vs. defender = 1/900

Phisher's strategy C2:

Imperfect Detection (exogenous)



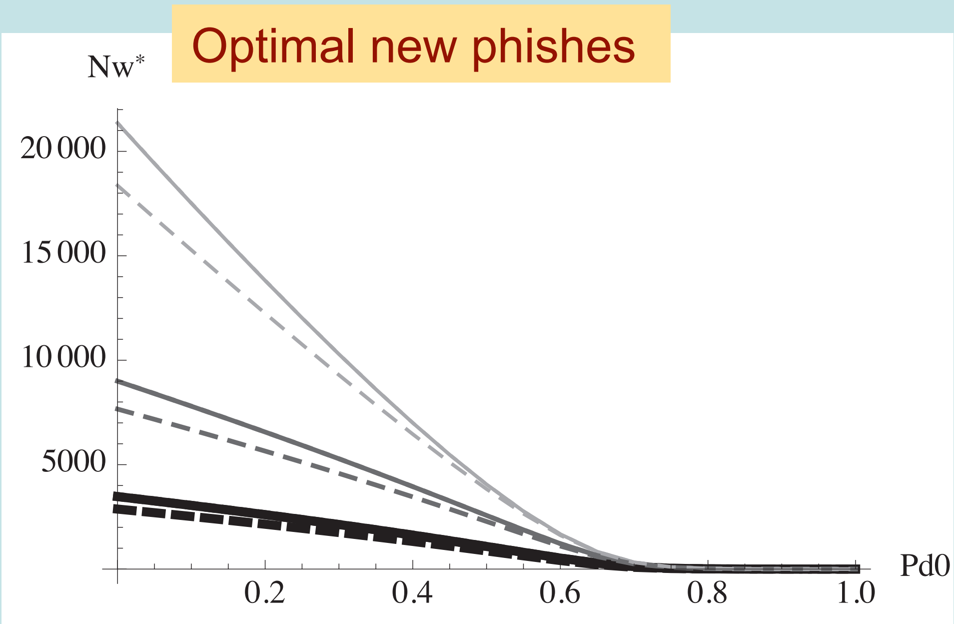
- Weak attacker creates more new phishes
- Weak attacker hurts more as Pd increases

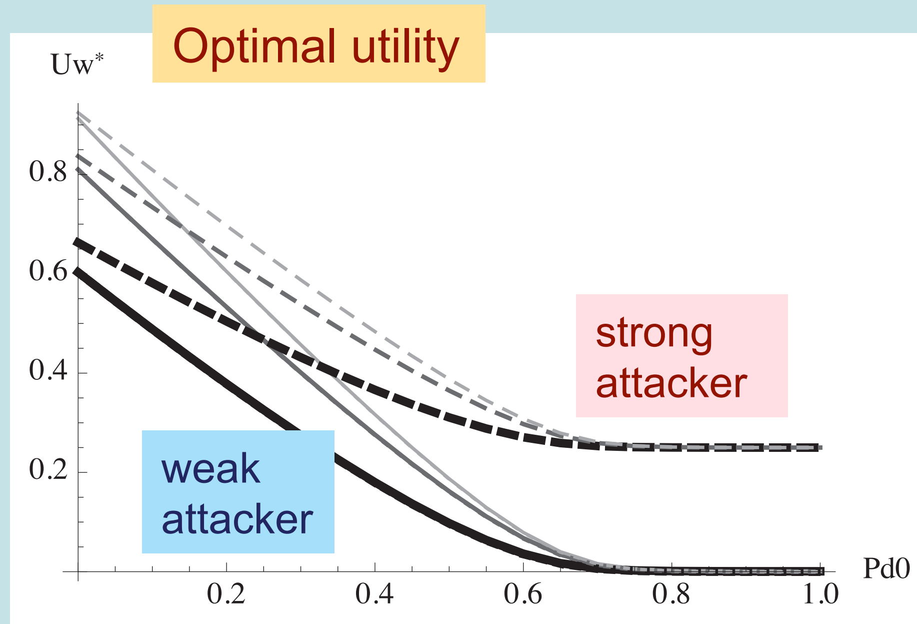
  better off, if Pd ➔ 1: improve resources to resist takedown

  if Pd ➔ 0: lower cost to create more phishes

Phisher's strategy C3:

Imperfect Detection (endogenous)





- If new phishes increase detection rate
  - Registrars look for suspicious domain registration pattern [6]
  - 'Rock Phish' and 'Avalanche' phishes hosted on same domain [APWG]
- Less phishes and utility

Discussion & Summary

Implications to Anti-Phishing Industry

- Increasing cost of a phish
  - Affect a weak attacker more
  - But can use stolen credit cards, or ‘easy’ domains (e.g., .tk, co.cc) [6]
  - 80% attacks used compromised servers [6,7]
- Improving detection rate
  - Concerns for sharing among takedown companies
  - User reporting (not necessarily requiring user evaluation) can be helpful
- Empirical estimation & prioritizing
  - Pd ➔ 0: make phishing cost higher
  - Pd ➔ 1: disrupt resources (e.g., access to botnet, underground market)

Summary

- Colonel Blotto Phishing (CBP)
  - Resource asymmetry
  - Information asymmetry
  - Endogenous dimensionality
- Applicability to web security problems
  - Two-step detect & takedown process
- Extensions
  - Competition between phishers -- Tragedy of the Commons? [8]

Reference

1. E. Borel. La theorie du jeu les equations integrales a noyau symetrique. Comptes Rendus de l'Academie des Sciences, 173:1304–1308, 1921.
2. E. Borel and J. Ville. Application de la theorie des probabilities aux jeux de hasard. Paris: Gauthier-Villars 1938.
3. O. A. Gross and R. A. Wagner. A continuous colonel blotto game. RAND Corporation RM–408, 1950.
4. B. Roberson. The colonel blotto game. Economic Theory, 29(1):1–24, Sept. 2006.
5. D. Kovenock, M. J. Mauboussin, and B. Roberson. Asymmetric conflicts with endogenous dimensionality. Purdue University Economics Working Papers 1259, Dec. 2010.
6. APWG. Global phishing survey: Trends and domain name use in 2H2010.
7. T. Moore and R. Clayton. Evil searching: Compromise and recompromise of internet hosts for phishing. In *FC* 2009.
8. C. Herley and D. Florencio. A profitless endeavor: phishing as tragedy of the commons. In *NSPW* 2008.

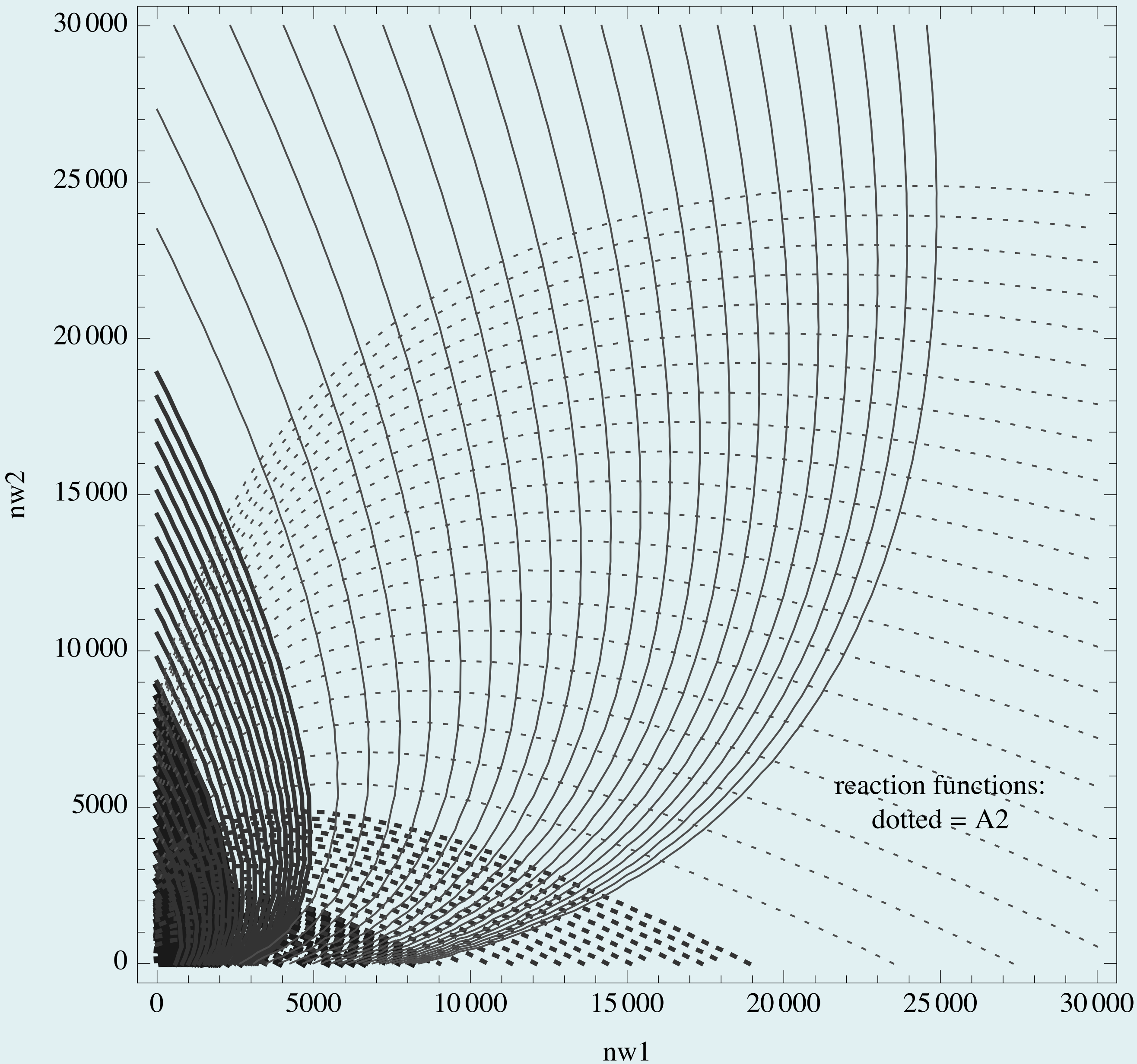
Thank you. Questions?

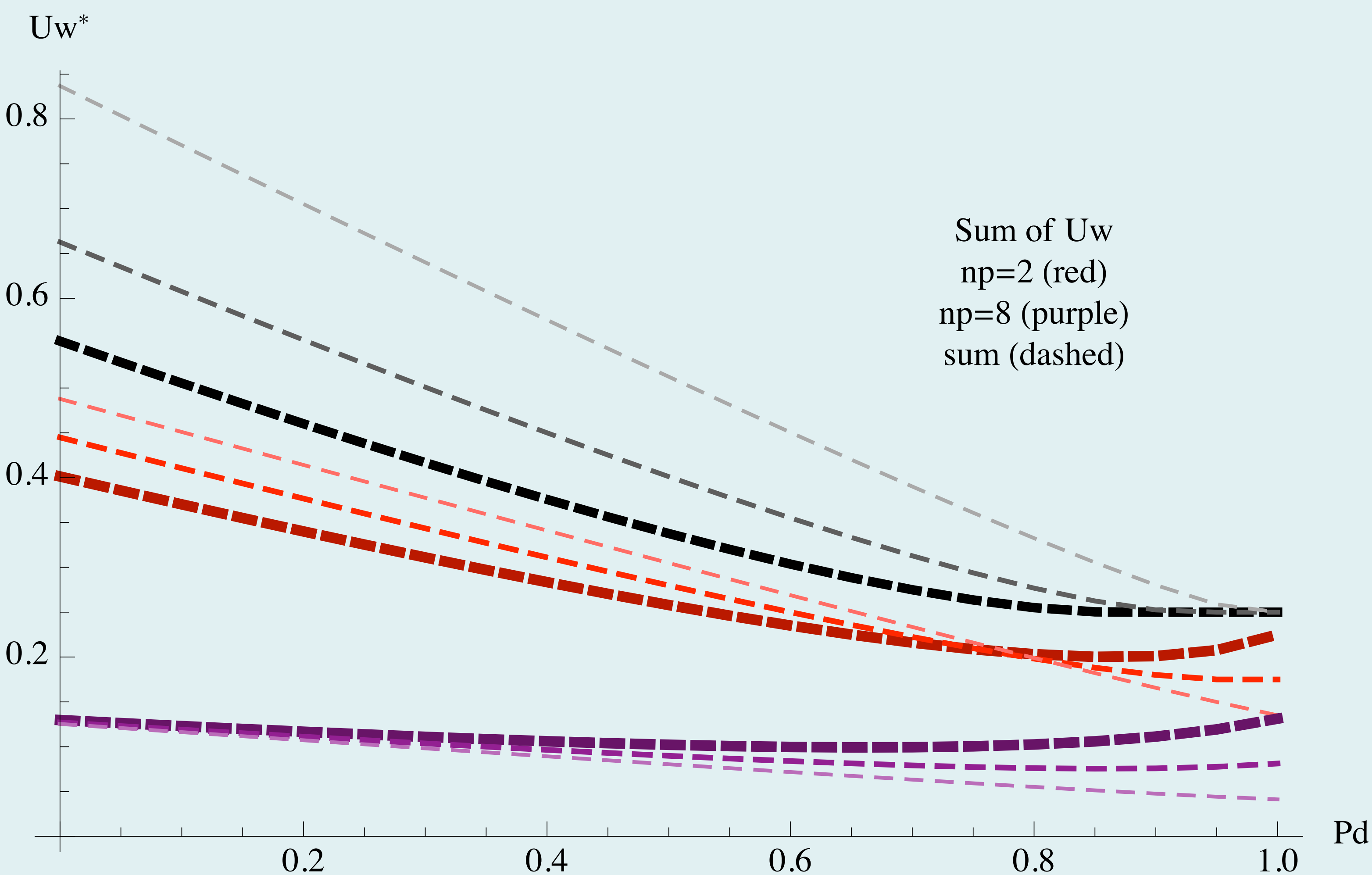
Pern Hui Chia

chia@q2s.ntnu.no

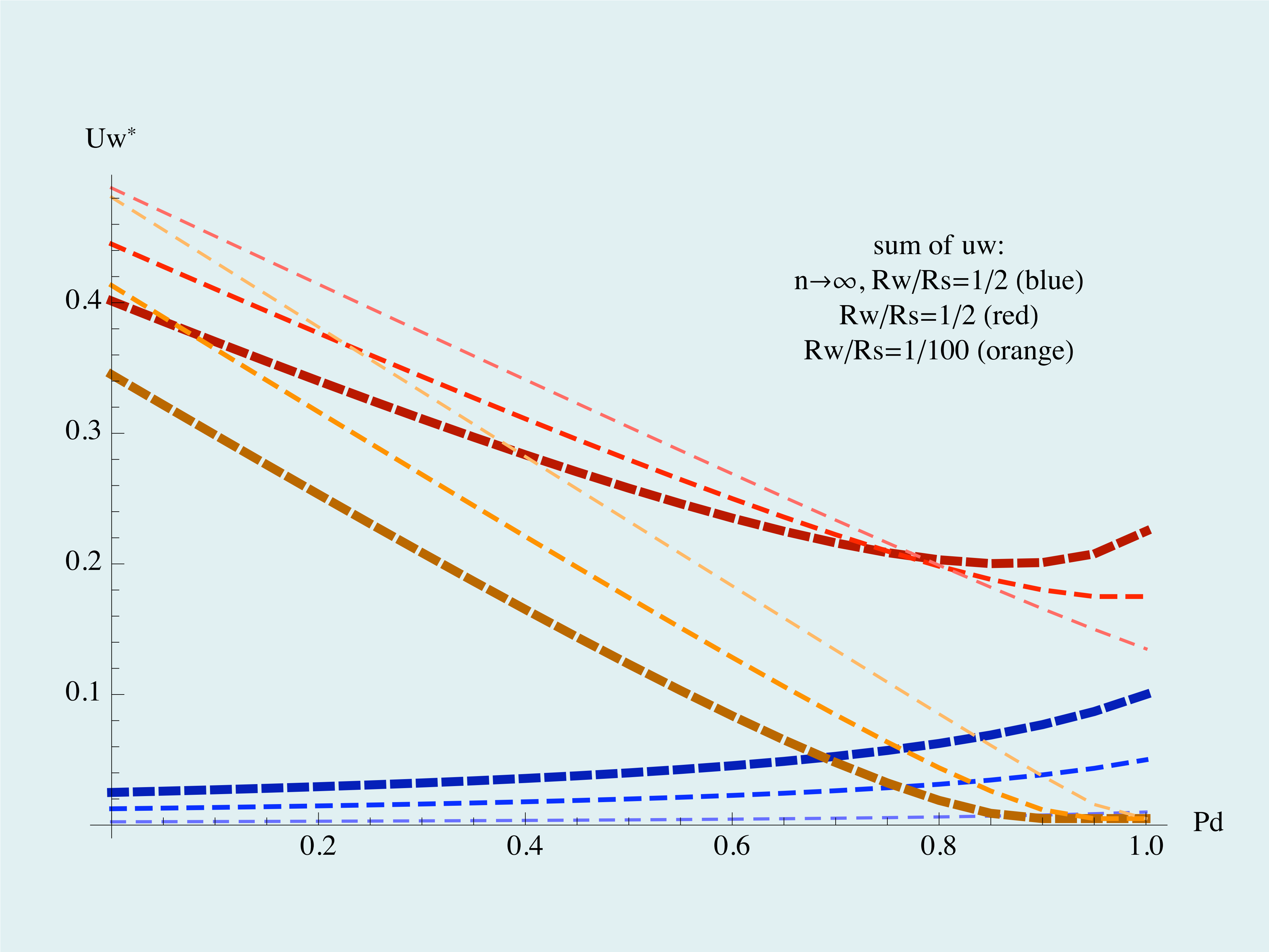
John Chuang

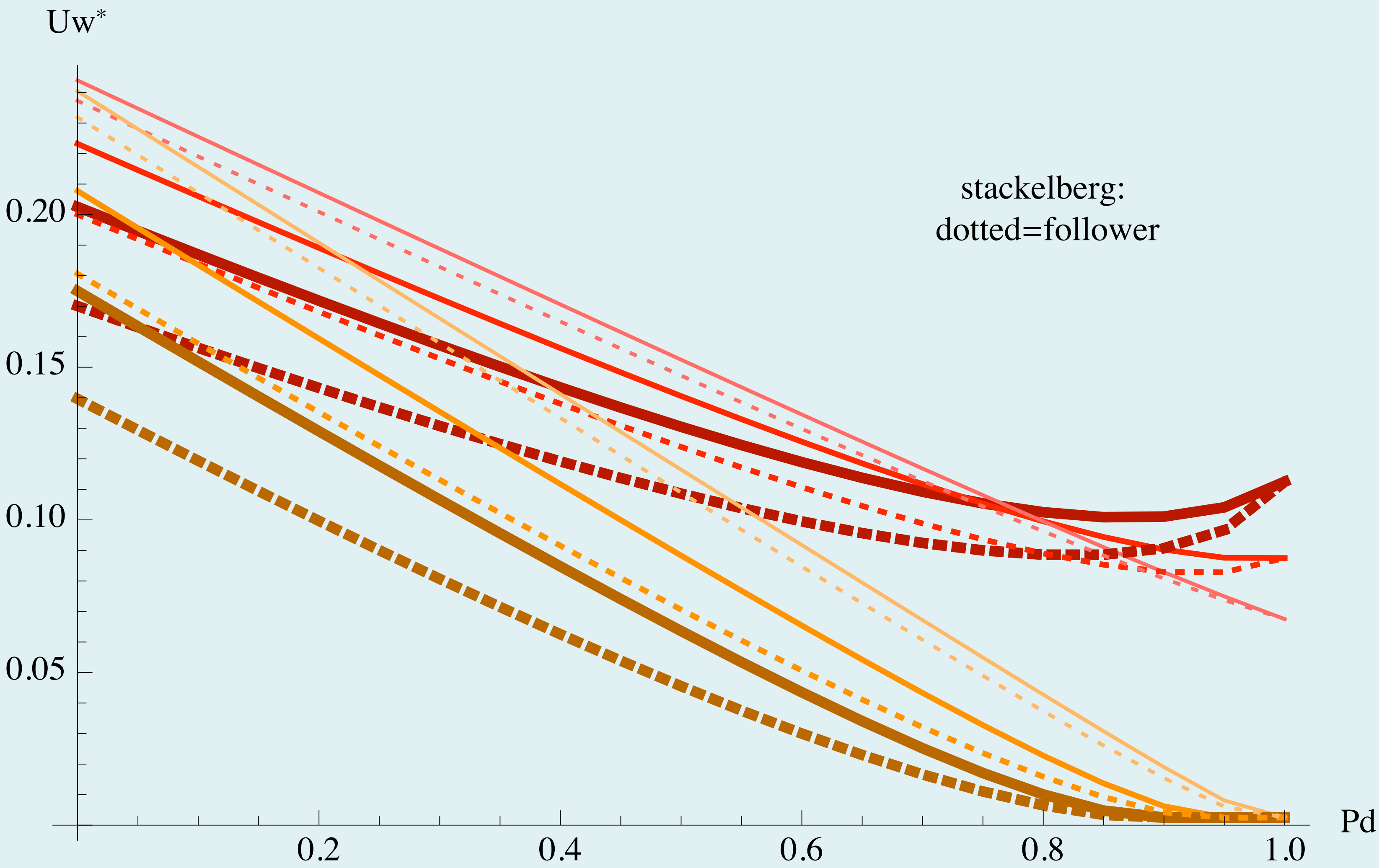
chuang@ischool.berkeley.edu


30 000
25 000
20 000
15 000
10 000
5000
0
nw2
0
5000
10 000
15 000
20 000
25 000
30 000
nw1
reaction functions:
dotted = A2



Uw*
0.8
0.6
0.4
0.2
Sum of Uw
np=2 (red)
np=8 (purple)
sum (dashed)
0.2
0.4
0.6
0.8
1.0
Pd


Uw*
sum of uw:
n→∞, Rw/Rs=1/2 (blue)
Rw/Rs=1/2 (red)
Rw/Rs=1/100 (orange)
0.4
0.3
0.2
0.1
0.2
0.4
0.6
0.8
1.0
Pd



Uw*
0.20
0.15
0.10
0.05
0.2
0.4
0.6
0.8
1.0
Pd
stackelberg:
dotted=follower