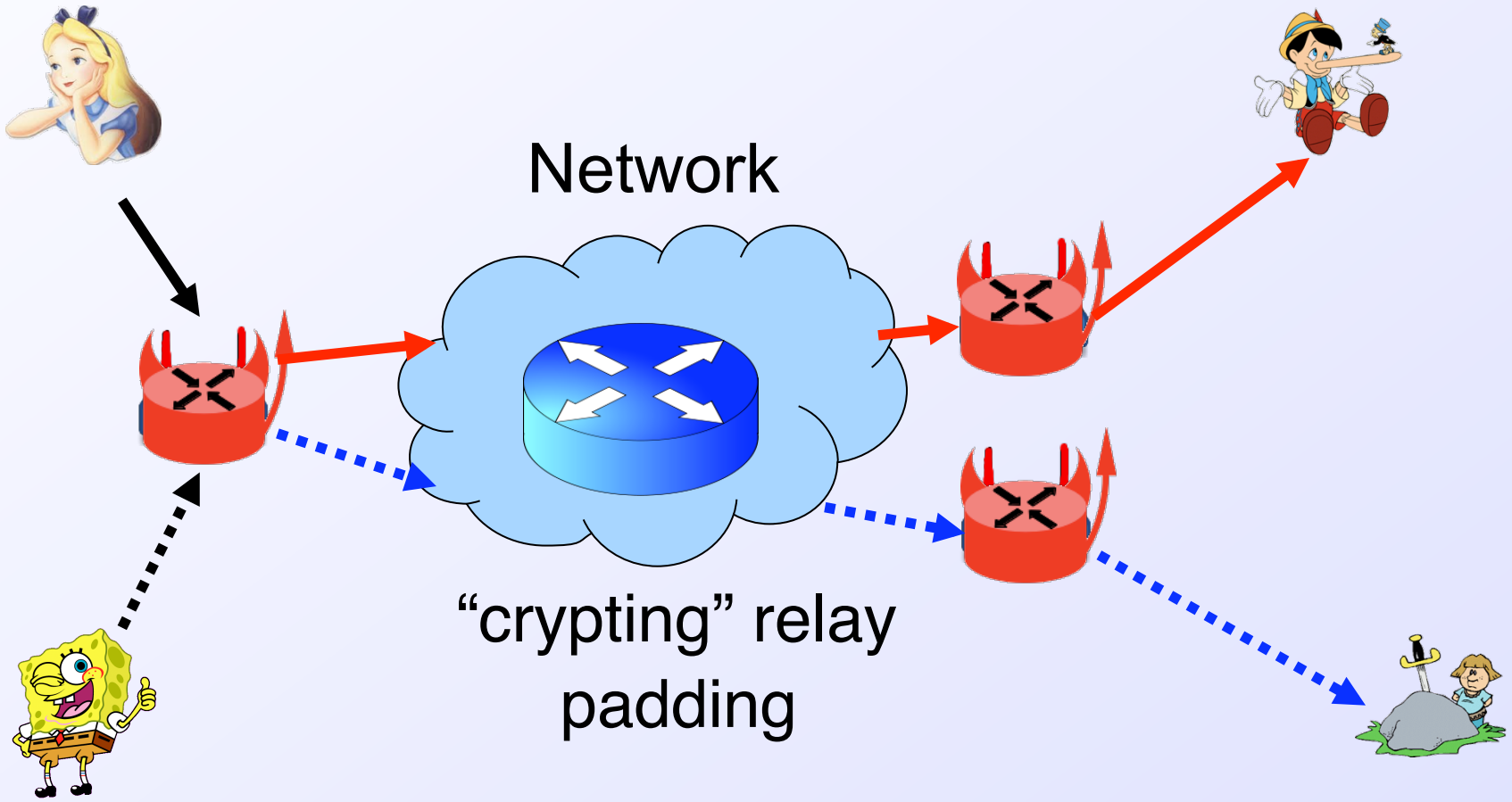
NEW ATTACKS ON NETWORK FLOW WATERMARKS

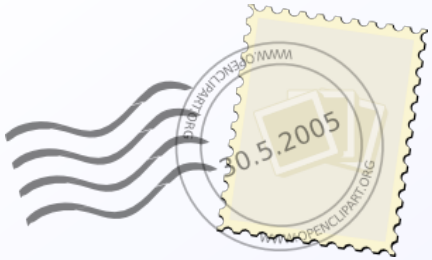
ZI LIN
NICK HOPPER
UNIVERSITY OF MINNESOTA

NETWORK FLOW WATERMARKING



Solution: "mark" packet delivery times

REQUIREMENTS



Watermark should be **hard to remove** and **unlikely to occur**



Robust to “natural” network jitter



Adversarial timing “noise” seems hard.



Invisible to prevent detection

PASSIVE INVISIBILITY





Warden





Flow(s) chosen from some “typical” distribution.

Warden sees flow(s) after marking/jitter.

Warden guesses “marked” or “unmarked”

ACTIVE INVISIBILITY





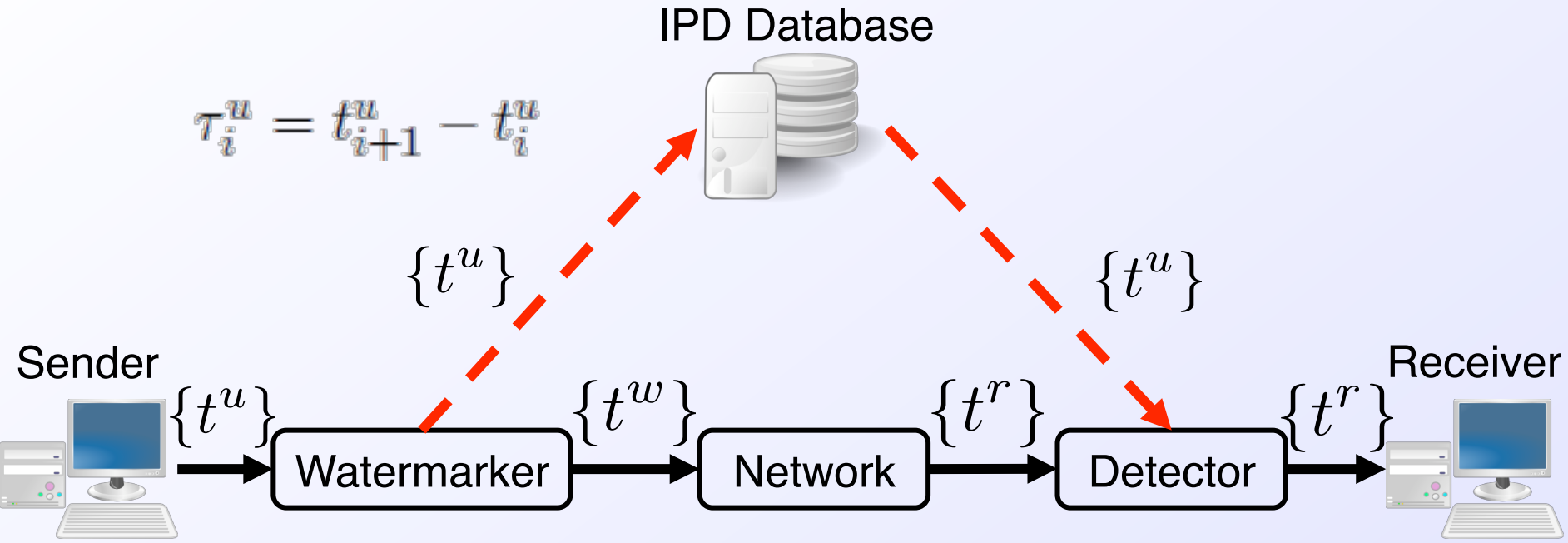




Warden chooses flow(s) **before** marking

Warden sees flow(s) after marking/jitter

RAINBOW*



$$\tau_i^w = \tau_i^u + w_i$$

$$w_i = \begin{cases} +a & \text{w. p. } \frac{1}{2} \\ -a & \text{w. p. } \frac{1}{2} \end{cases}$$

$$\tau_i^r = t_{i+1}^r - t_i^r = \tau_i^u + w_i + \delta_i$$

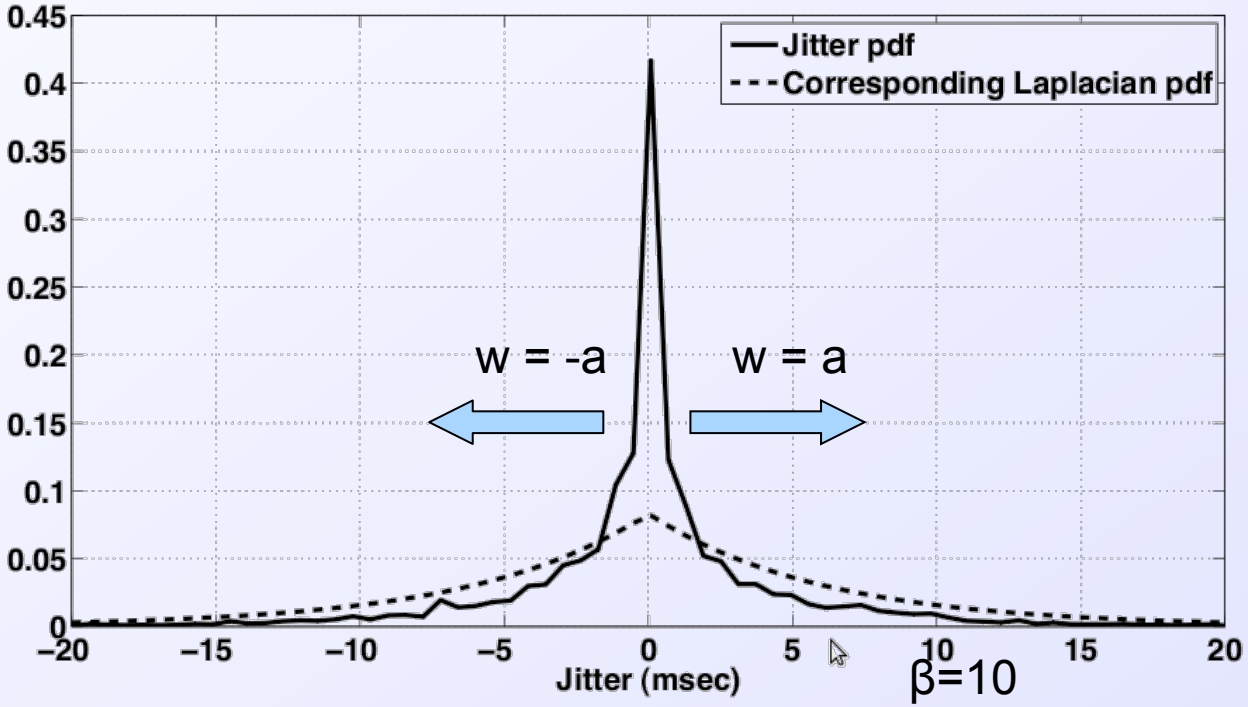
marked if

$$\langle \tau^r - \tau^u, w \rangle \geq \frac{a}{2\sigma_{\text{jitter}}}$$

**A. Houmansadr, N. Kiyavash, and N. Borisov. "RAINBOW: A robust and invisible non-blind watermark for network flows." NDSS 2009.*

DETECTING RAINBOW

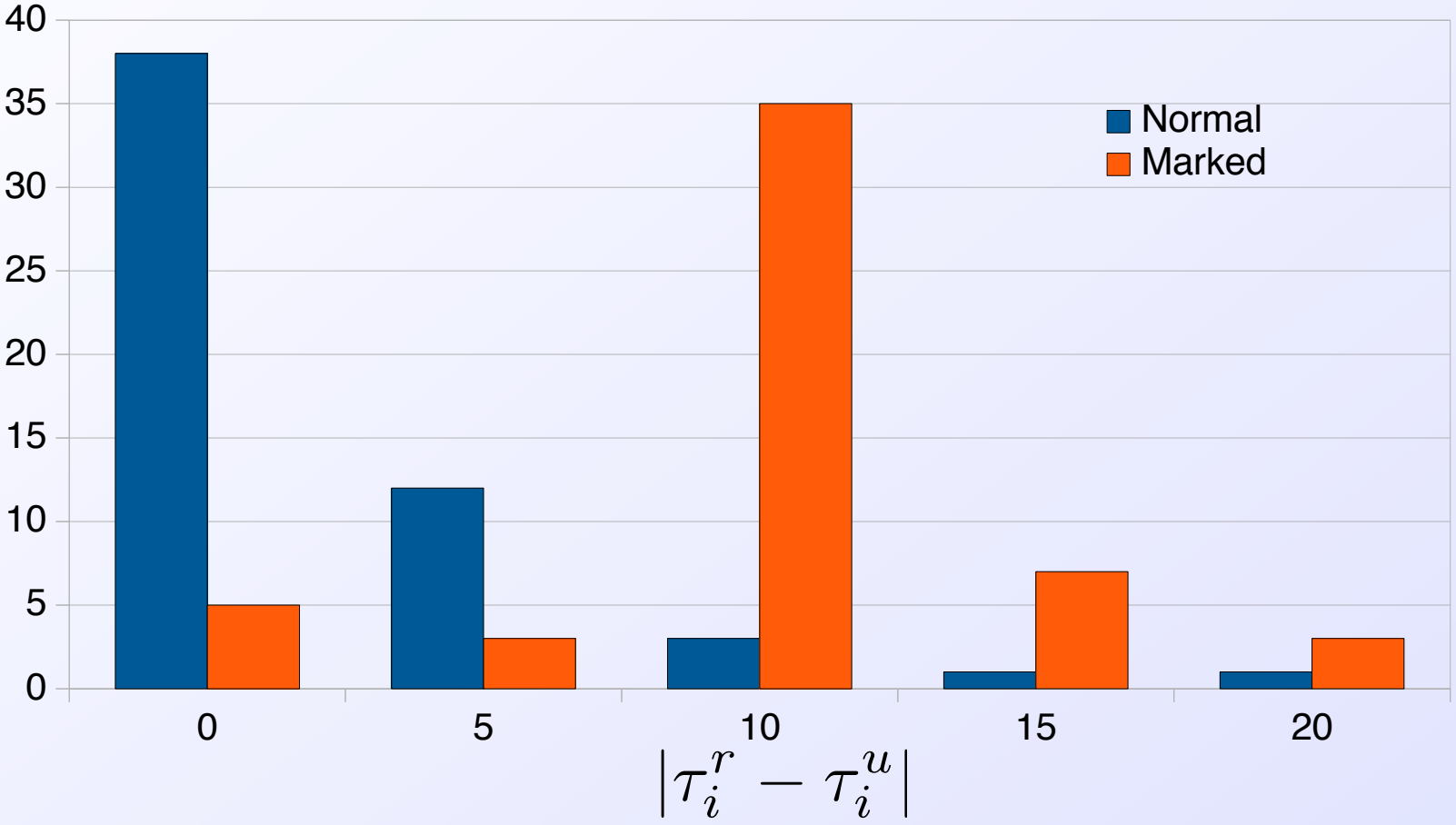
$$\tau_i^r = t_{i+1}^r - t_i^r = \tau_i^u + w_i + \delta_i$$



β=10

KNOWN FLOW ATTACK

Use known IPDs (τ^u) to compute jitter histogram:

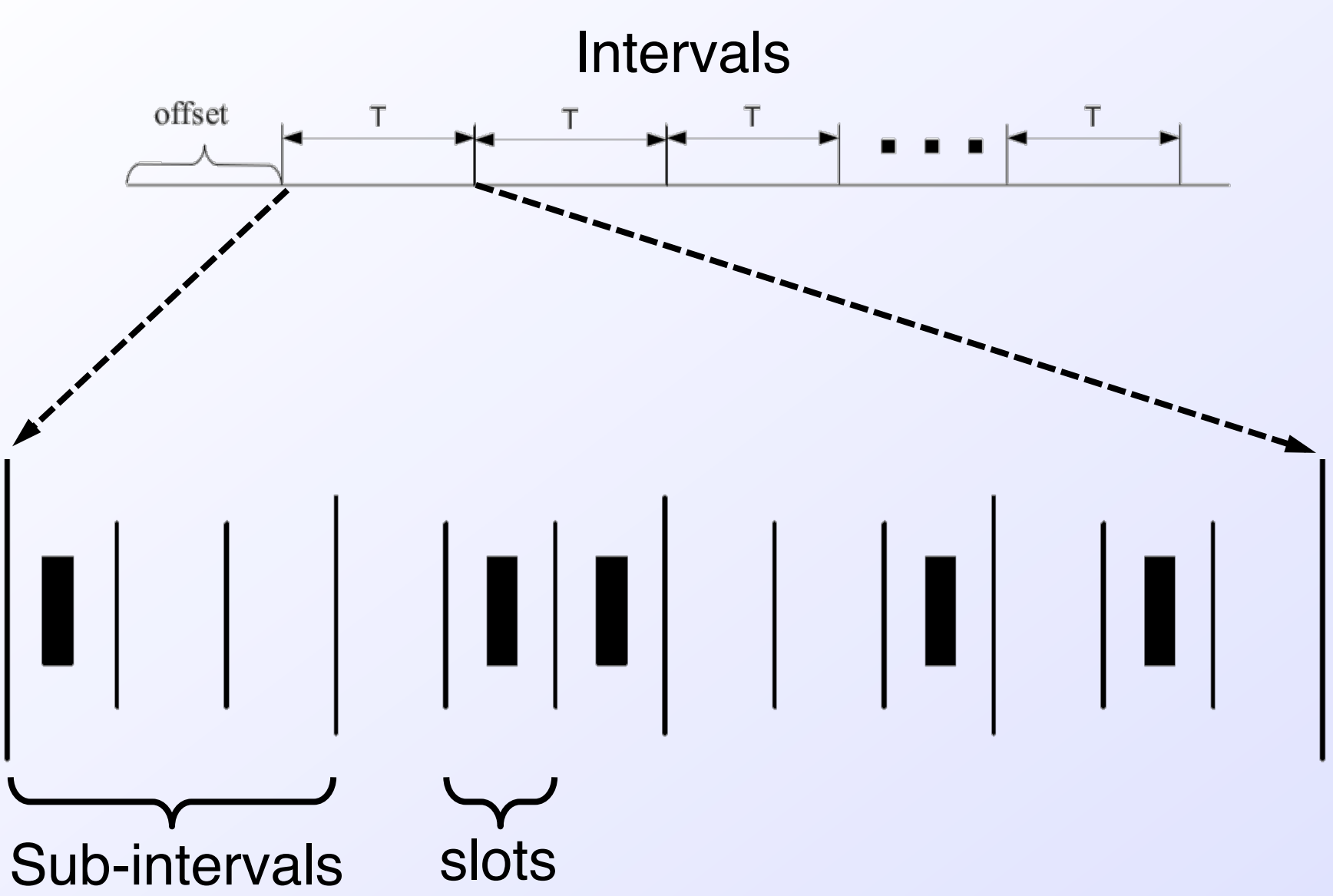


Marked if Pr[$|\delta_i| < \beta/4$] < $(1-e^{-1/4})/2 \approx 0.11$

RESULTS

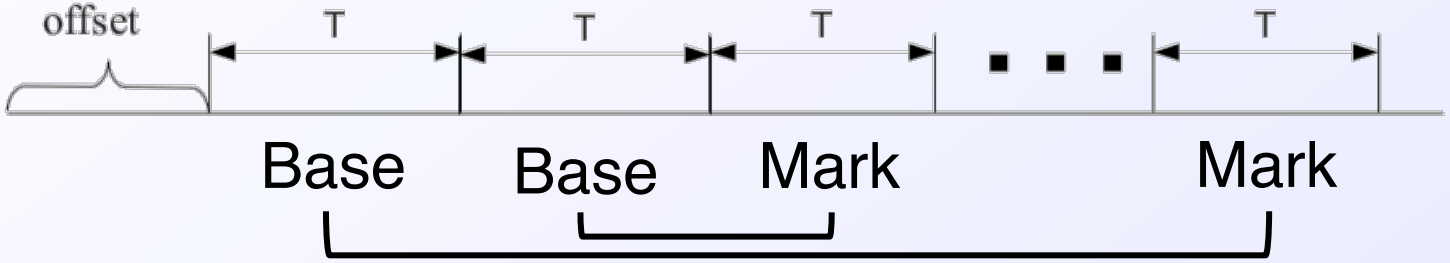
a (ms)	length	TPR	FPR	Avg. Recall
2	200	0.004	0	0.001
	500	0.011	0	0.011
	1000	0.024	0	0.001
5	200	0.996	0	0.692
	500	1.000	0	0.846
	1000	1.000	0	0.921
10	200	1.000	0	0.984
	500	1.000	0	0.994
	1000	1.000	0	0.997
15	200	1.000	0	1.000
	500	1.000	0	1.000
	1000	1.000	0	1.000
20	200	1.000	0	1.000
	500	1.000	0	1.000
	1000	1.000	0	1.000

SWIRL*

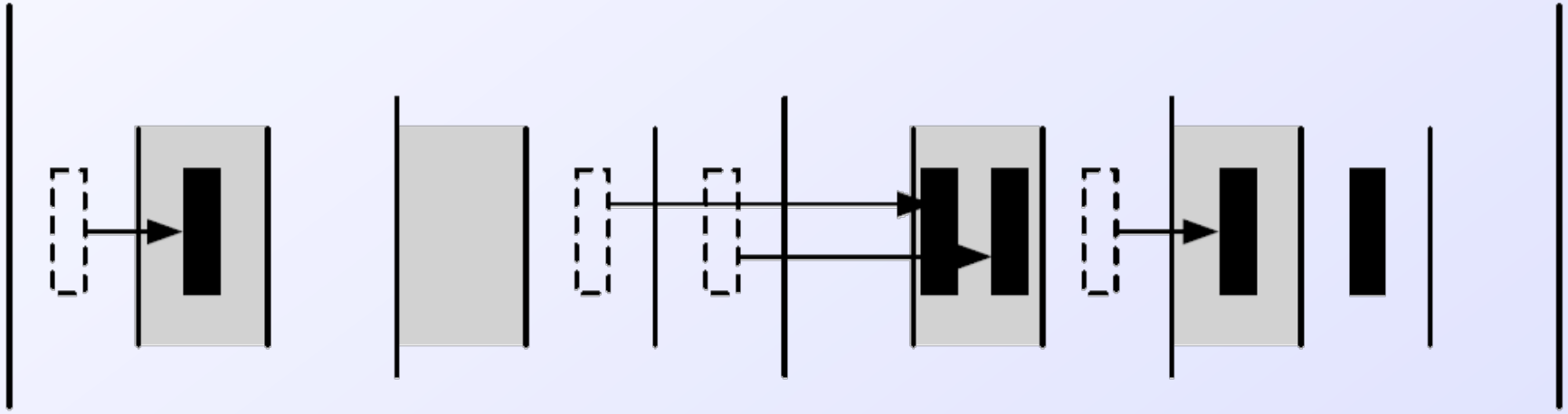


*A. Houmansadr, N. Borisov. "SWIRL: A scalable watermark to detect correlated network flows." NDSS 2011.

Base intervals determine *designated slots* in corresponding *marked* intervals:



Packets delayed to next designated slot:



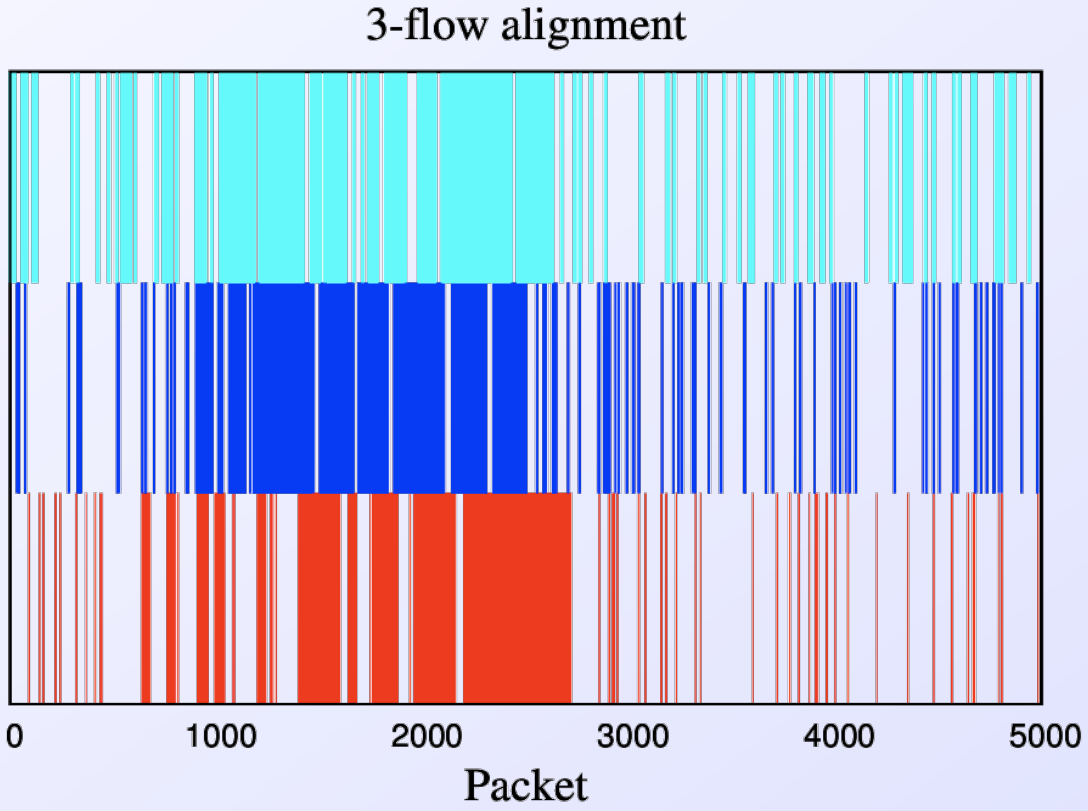
Marked if enough designated slots have packets.

DETECTING SWIRL

Chosen multi-flow: uniform k packets/sec

Flag packets “too close” or “too far” from previous packet

Simultaneous flags across flows indicate marked intervals

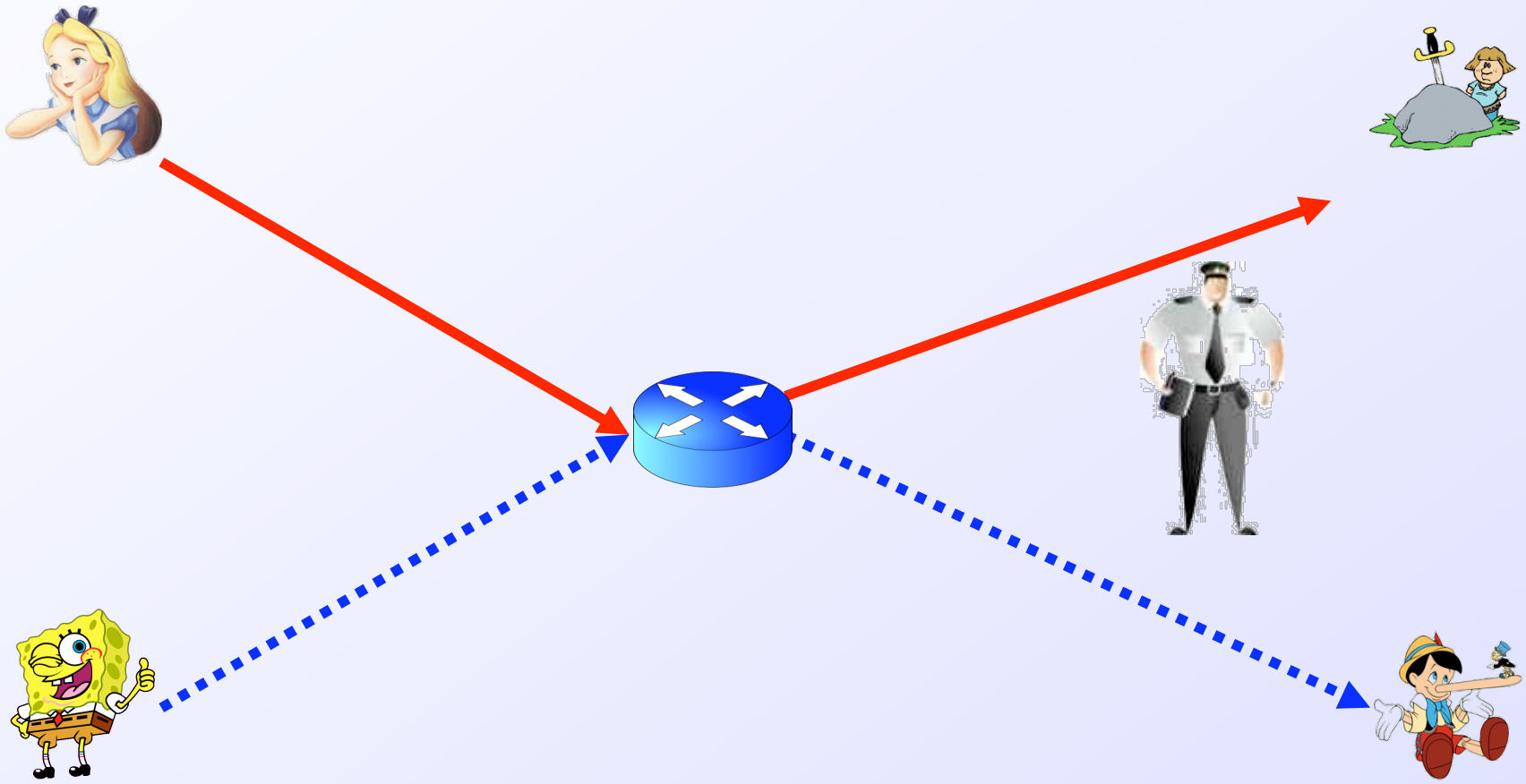


Simulation: 100% accuracy across 100 experiments

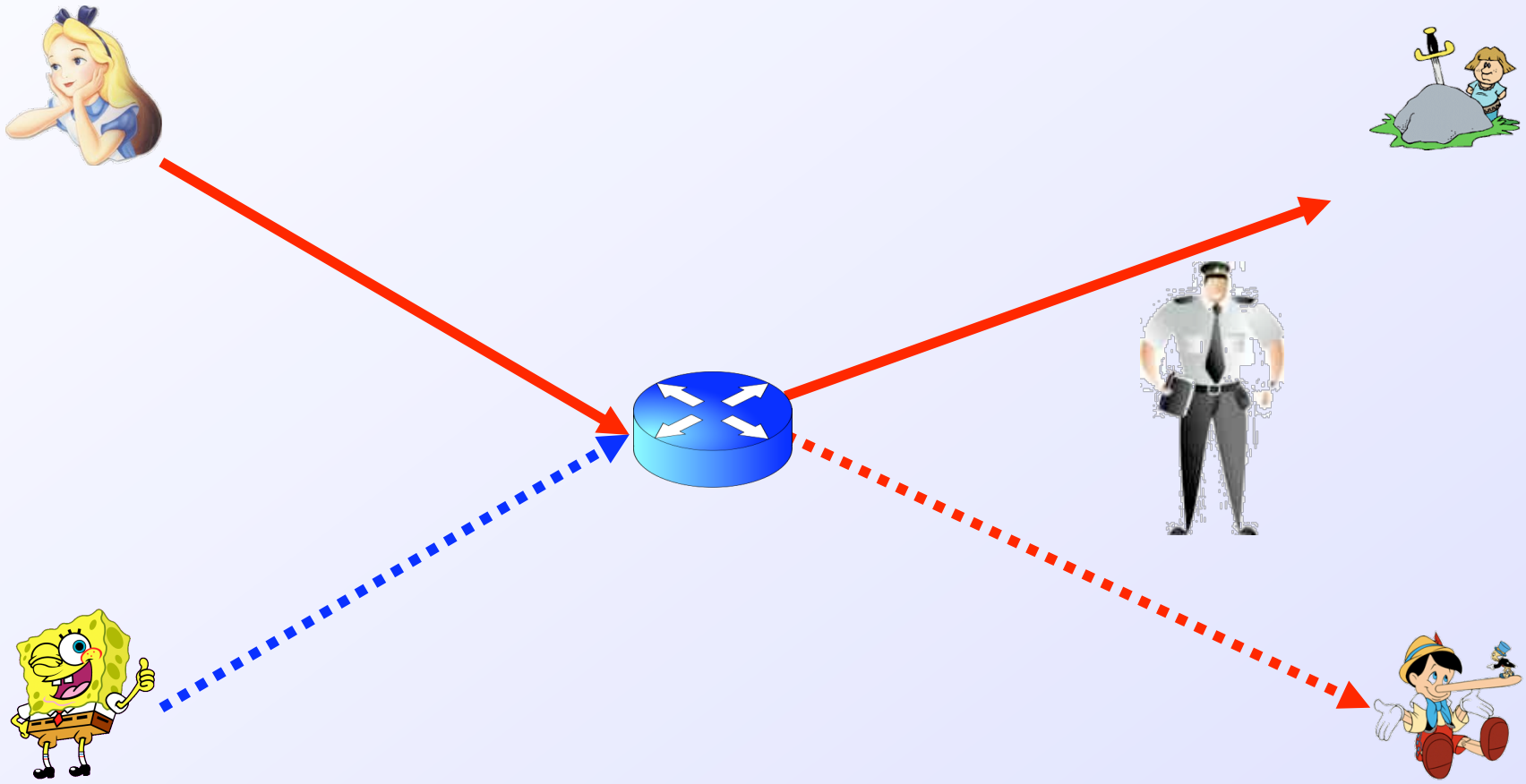
ACTIVE ATTACKS



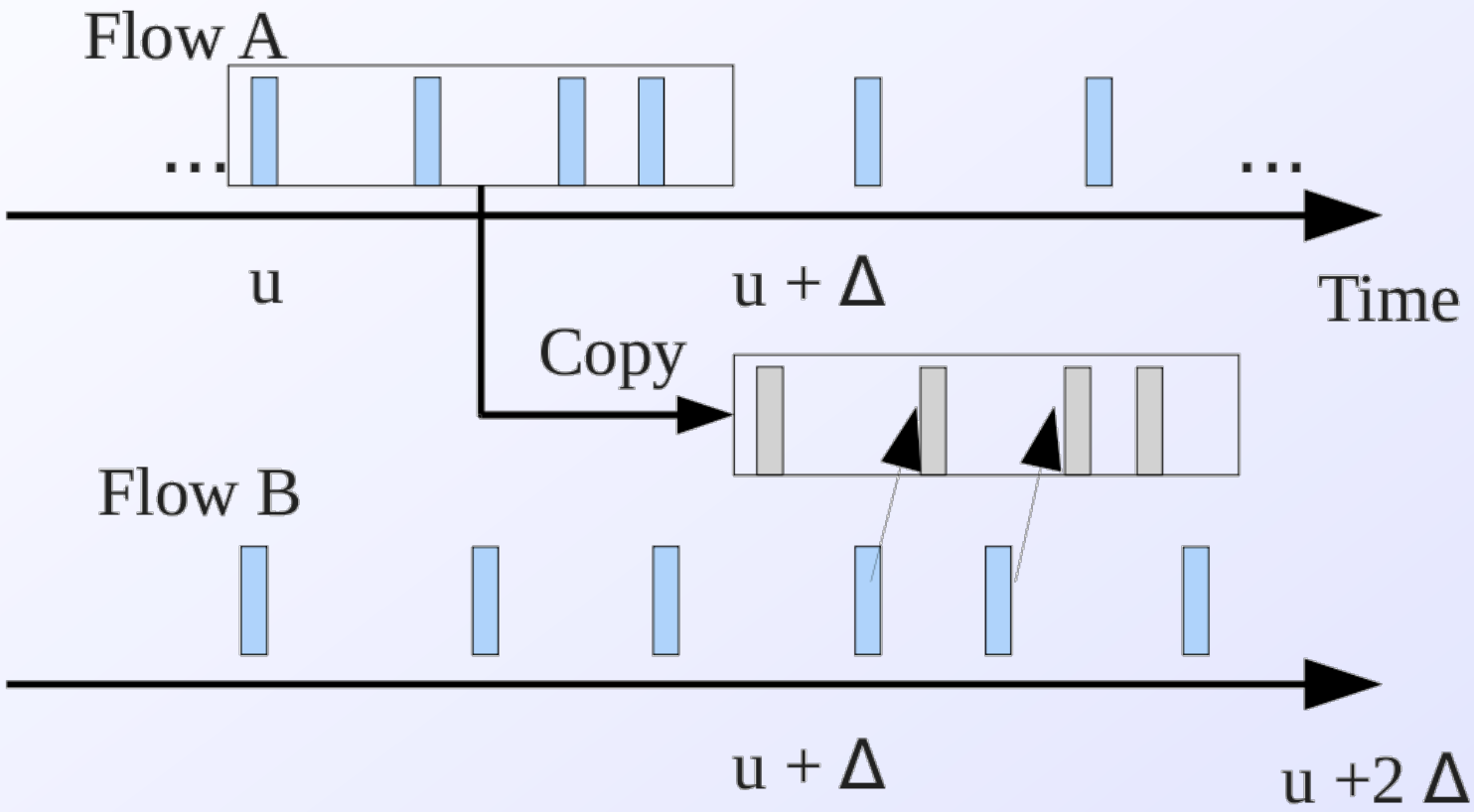
ACTIVE ATTACKS



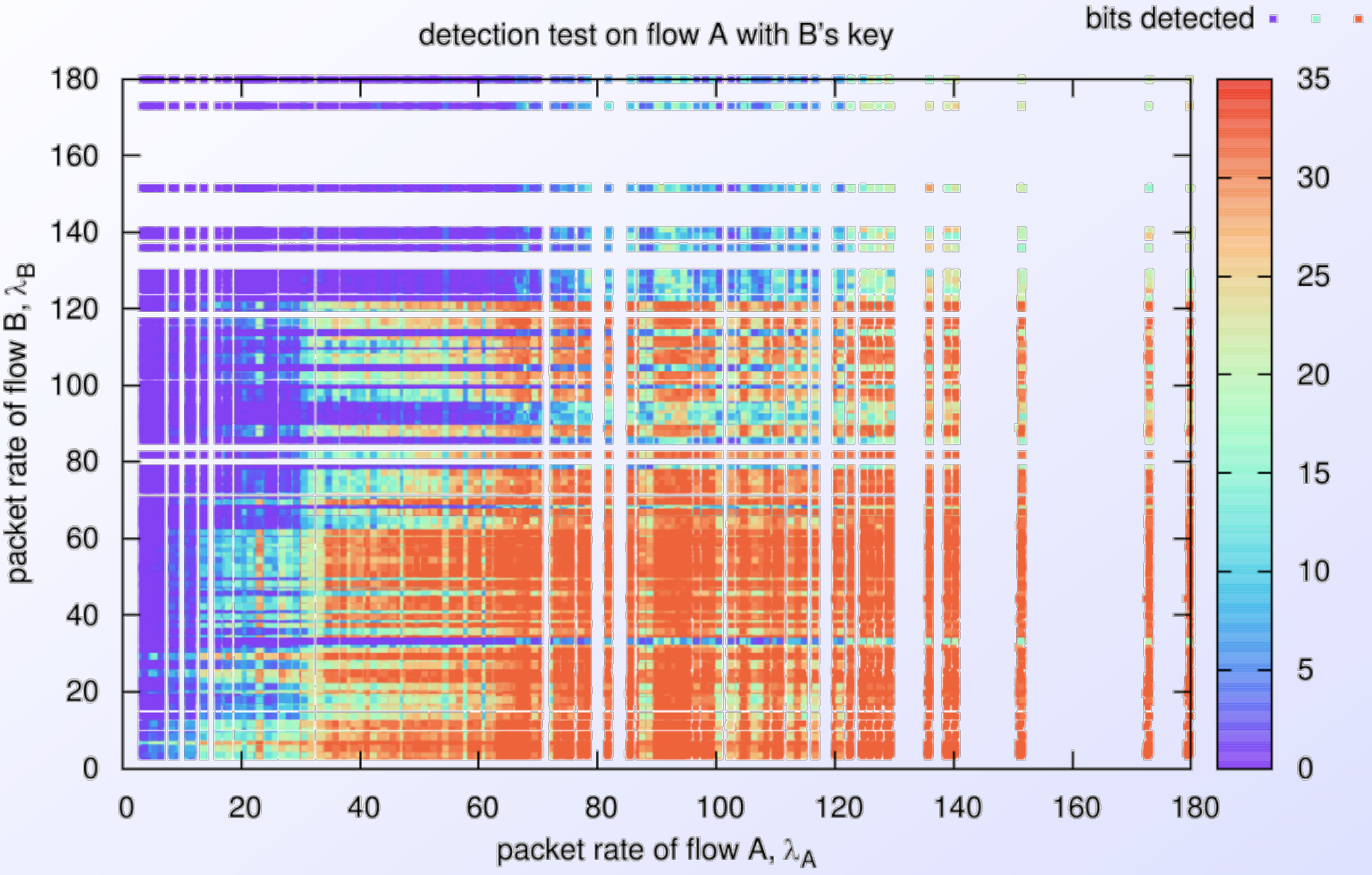
COPY ATTACK



COPYING SWIRL



RESULTS OF COPY ATTACK



THANK YOU