Feedback

(1) in 5th page. See the highlight part below, it lose a ")".

set the entries $\{p^k_{i_{kl}}\}_{kl \in I^V_{max}} = 1$ where I^V_{max} indicates the set of kl^{th} entries with highest velocities in the matrix V^k_i (i.e. $(\sim, I^V_{max} = sort(abs(V^k_i(:))))$, and 2) $\{p_{kl}\}$ is set to one with probability $sigmoid(v_{ij})$, where $sigmoid(x) := \frac{1}{1+e^{-x}}$, otherwise it is set to zero. The PSO process is continued for N_i iterations.

$$V^k_i = \omega V^{k-1}_i + \alpha_1 \beta_1 (BP^k_i - P^{k-1}_i) + \alpha_2 \beta_2 (GP^k - P^{k-1}_i) \quad (4)$$

(2) in 5th page. See the highlight part below. r_c should be r.

pair of parents to generate new children (offsprings). Eq.(3) defines the crossover operation where r_c denotes a randomly chosen row from set $\{1, ..., n\}$. These Offsprings form part of the new chromosomes of the next generation. To increase the diversity of the population, the mutation operation is performed on each child where the mutation operator changes an entry of sampling matrix C from zero-to-one or vice-versa with probability p_m. The GA process is continued over N_i iterations and the best chromosome in each iteration remains unchanged. In most cases throughout this paper, the GA parameters are set as: $N_i = 60$, $N_p = 1500$, $p_c = 0.3$, and $p_m = 0.01$.

$$\text{OffSpring}_1 = C_1(1:r,:) + C_2(r+1:n,:)$$
$$\text{OffSpring}_2 = C_2(1:r,:) + C_1(r+1:n,:) \quad (3)$$

(3) in 7th page.

SR	0.0088	0.0265	0.0442	0.0619	0.0796
P^d_{CP}	0.8077	0.9135	0.9306	0.9499	0.9566
P^{fa}_{CP}	0.4749	0.2412	0.1661	0.1263	0.1032

TABLE II: Average P^d_{CP} and P^{fa}_{CP} for Harvard network.

The results in the above table is calculated as follows: (This method is used at an early time)

1) firstly, construct a binary delay matrix X according to the real delay matrix;
2) then use X and MC to obtain an estimated delay matrix Y;
3) change Y to be a binary matrix;
4) finally, compare two binary matrices X, Y to calculate the probability pd and pfa.

The results below is firstly to obtain an estimated matrix Xhat according the real delay matrix X.(The method is used in calculating NMSE) And let the threshold to be the

average delay in real delay matrix. Then use the threshold to compare X and Xhat to calculate pd and pfa.

	$P^d{}_{CP}$ Harvard(GA)	$P^{fa}{}_{CP}$ Harvard(GA)	$P^d{}_{CP}$ Harvard(RS)	$P^{fa}{}_{CP}$ Harvard(RS)
SR = 0.0088	0.014053	0.003889	0	0
SR = 0.0177	0.60802	0.162028	0.429819	0.130273
SR = 0.265	0.660917	0.144711	0.550066	0.132213
SR = 0.0354	0.753422	0.126221	0.635393	0.128989
SR = 0.0443	0.790478	0.108758	0.722461	0.108994
SR = 0.0531	0.835796	0.074993	0.799928	0.091303
SR = 0.0619	0.869354	0.061548	0.852101	0.07899
SR = 0.0708	0.906074	0.053057	0.887573	0.066636
SR = 0.0796	0.925804	0.045195	0.913601	0.05692
SR = 0.0885	0.937756	0.039888	0.93153	0.048441

(4) in 7th page. See the highlight part below. PG := 100% × .....

for per-flow size estimation, representing the fact that in low sampling rates the intelligent design of the observation matrix using PSO algorithm results in a better estimation accuracy. The reduction in the computational complexity using the PSO algorithm is quantified using the notion of Processing Gain (PG) defined as $PG := 100 \times \frac{PT_{GA} - PT_{PSO}}{PT_{GA}}$ where PT_{GA} and PT_{PSO} respectively denote the processing times for running GA and PSO algorithms. Accordingly, the processing gains for Abilene and Geant networks are PG=56% and PG=65%, respectively.

E. Deployability of SNIPER in Dynamic Environments

(5) in 8th page. The data of table IV and pd, pfa is sended to Prof wang. Because the data have two types(k% or k largest flows in test), there does not list the results.