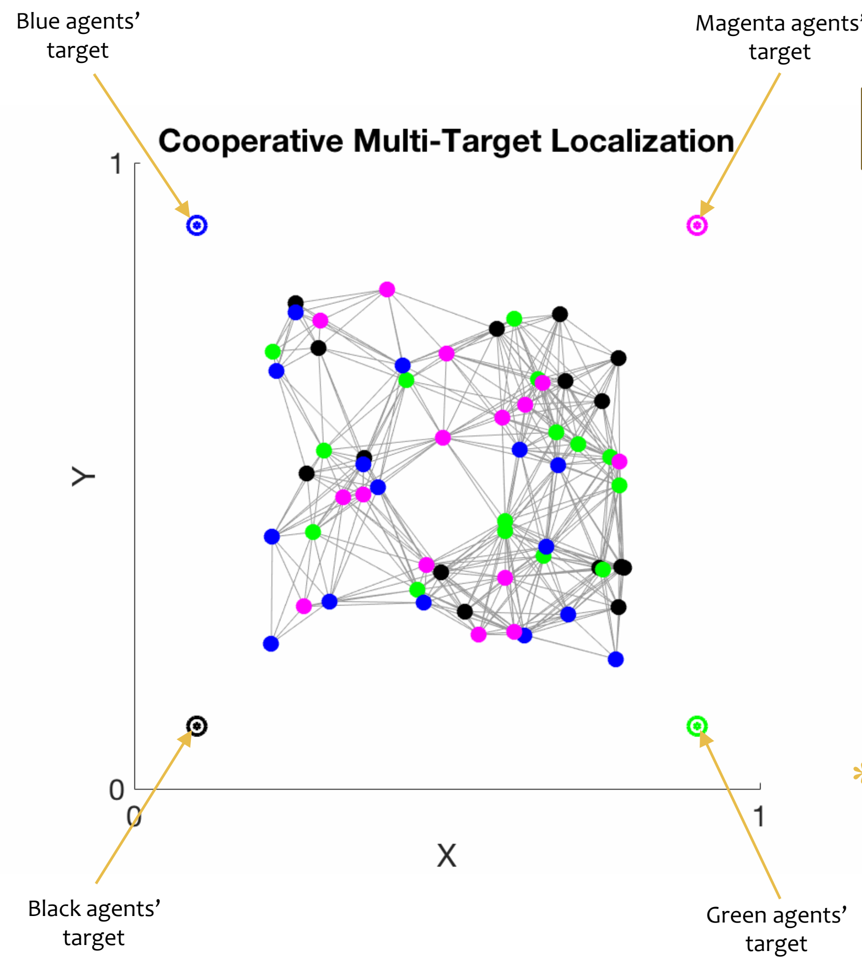
Resilient Diffusion Least-Mean Squares over Adaptive Networks for Distributed Clustering in CPS

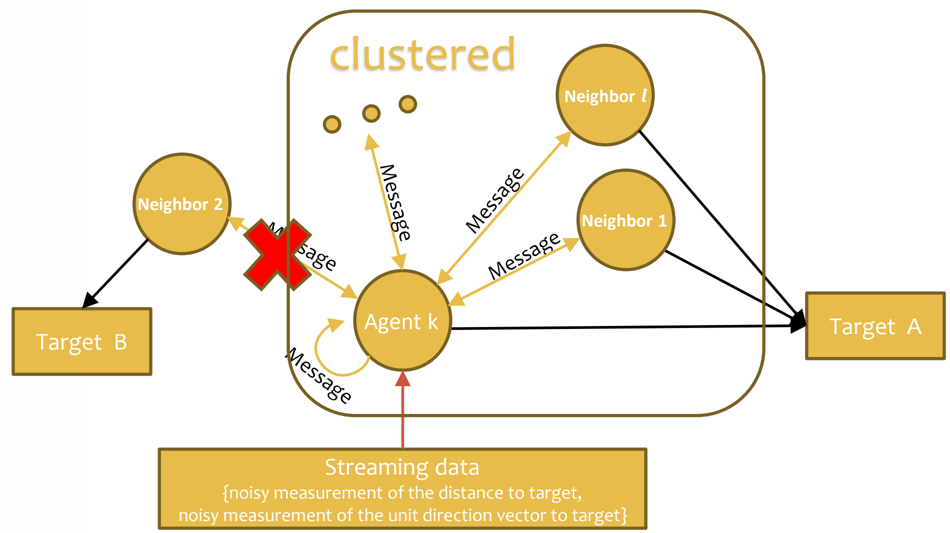
Jiani Li and Xenofon Koutsoukos

Vanderbilt University

Motivating Application:
Cooperative Multi-Target Localization



Agent k's learning and clustering procedure



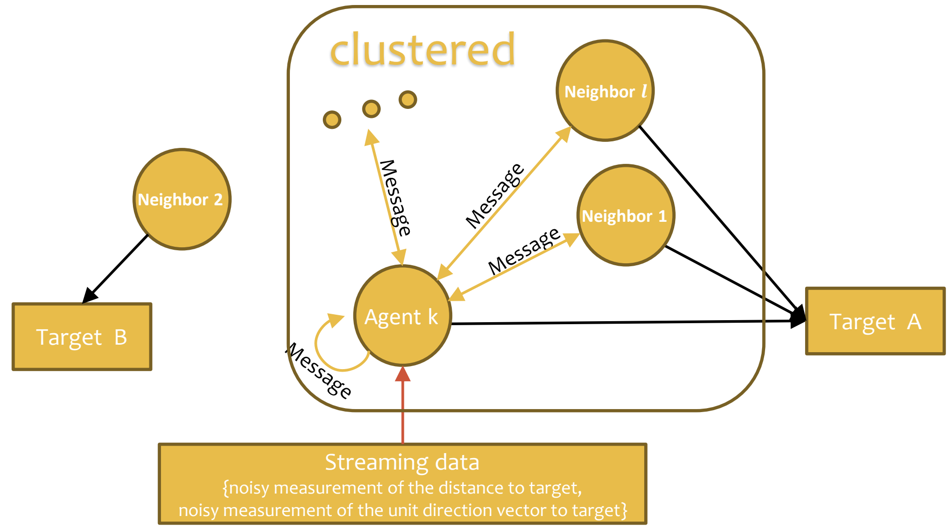
* Problem Formulation:
  * Distributed estimation
  * Multi-tasks network
  * Clustering for better estimation performance

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Motivating Application:
Cooperative Multi-Target Localization

* Other Applications:
  * Cooperative data mining
  * Multi-task learning
  * Distributed clustering
  * Intrusion detection
  * Static target localization
  * Real-time learning, adaptation
    * Mobile target localization
  * Spectrum sensing
  * Speech enhancement
  * Biological inspired design
    * Fish schooling
    * Bees swarming

Agent k's learning and clustering procedure



* Problem Formulation:
  * Distributed estimation (stationary/time-varying)
  * Multi-tasks network
  * Clustering for better estimation performance

Diffusion Least-Mean Squares over Adaptive Networks for Distributed Clustering

Algorithm 1 ATC diffusion strategy with adaptive combination weights

Set $\gamma^2_{lk}(-1) = 0$ for all $k = 1, 2, ..., N$ and $l \in N_k$

for all $k = 1, 2, ..., N, i \geq 0$ **do**

$e_k(i) = \boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}$

$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k \boldsymbol{u}^*_{k,i} e_k(i)$

$\gamma^2_{lk}(i) = (1-\nu_k)\gamma^2_{lk}(i-1) + \nu_k \|\boldsymbol{\psi}_{l,i} - \boldsymbol{w}_{k,i-1}\|^2, l \in N_k$

$a_{lk}(i) = \frac{\gamma^{-2}_{lk}(i)}{\sum_{m \in N_k} \gamma^{-2}_{mk}(i)}, l \in N_k$

$\boldsymbol{w}_{k,i} = \sum_{l \in N_k} a_{lk}(i)\boldsymbol{\psi}_{l,i}$

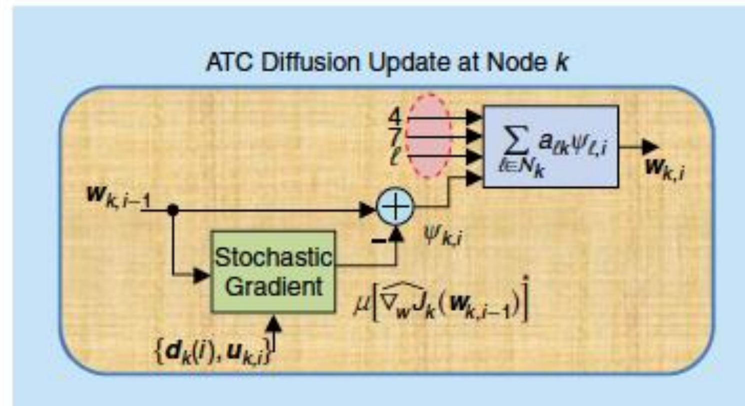
end for

adaptation

combination

Communication message

weight metrics



Agents assign large weights to neighbors estimating a similar model with its own.

Ali H. Sayed, Sheng-Yuan Tu, Jianshu Chen, Xiaochuan Zhao, Zaid J. Towfic: **Diffusion Strategies for Adaptation and Learning over Networks: An Examination of Distributed Strategies and Network Behavior.** IEEE Signal Process. Mag. 30(3): 155-171 (2013)

Diffusion Least-Mean Squares over Adaptive Networks for Distributed Clustering

Q: Are these algorithms resilient to cyber-attacks?

Attack Objectives

Assumption

Attacker knows the true model w_k^0

Drive normal agents to converge to a point as far from w_k^0 as possible.

Prolong the convergence time of the normal agents.

$$\max_{T_k} \|T_k\|$$

$$T \approx \frac{\ln\left(\frac{\epsilon N \cdot \mathrm{MSD}}{1 - N \cdot \mathrm{MSD}}\right)}{2 \ln\left(1 - \mu \mathrm{Tr}(R_u)/M\right)}$$

$$\max_{w_{k,T_k} \in [w_{\min}, w_{\max}]} \|w_k^0 - w_{k,T_k}\| \quad \begin{cases} \max_{w_k^a \in [w_{\min}, w_{\max}]} & \|w_k^0 - w_k^a\| \\ \min_{w_{k,T_k} \in [w_{\min}, w_{\max}]} & \|w_{k,T_k} - w_k^a\| \end{cases}$$

$$w_k^a = \begin{cases} w_{\min}, & \text{if } \|w_{\min} - w_k^0\| \geq \|w_{\max} - w_k^0\| \\ w_{\max}, & \text{if } \|w_{\min} - w_k^0\| < \|w_{\max} - w_k^0\| \end{cases}$$

$$\min_{w_{k,T_k} \in [w_{\min}, w_{\max}]} \quad \|w_{k,T_k} - w_k^a\|$$

Attacker does not know w_k^0

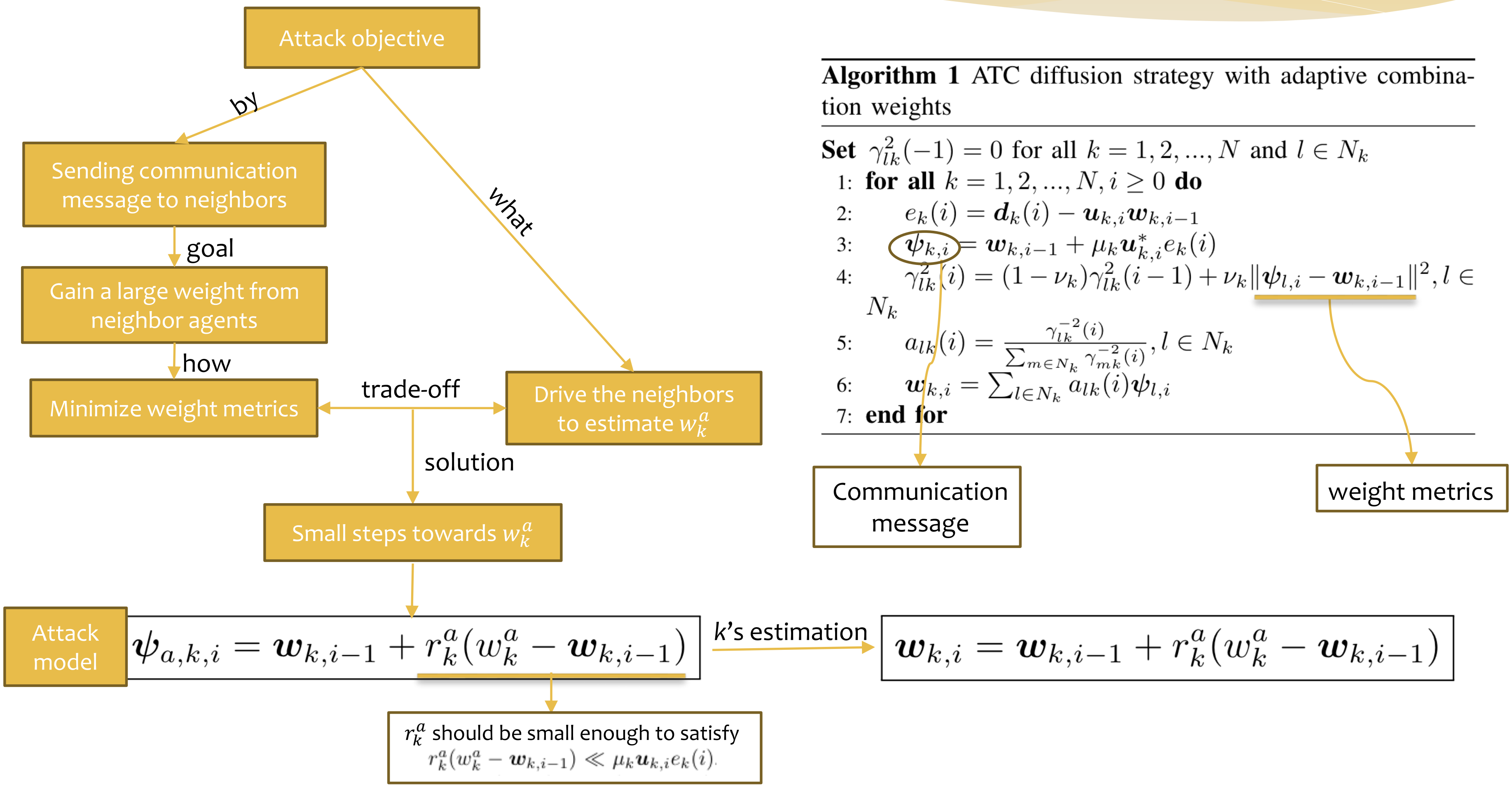
Drive normal agents to converge to a selected point w_k^a

$$\min_{w_{k,T_k} \in [w_{\min}, w_{\max}]} \quad \|w_{k,T_k} - w_k^a\|$$

- These objectives turn out to be represented in the same mathematical form.
- Under known w_k^0, attacker gets w_k^a by solving the maximization function.
- Under unknown w_k^0, attacker selects any w_k^a.
- For both known and unknown w_k^0, attacker needs to solve the minimization problem.
- That is, after entering stable state, attacker's neighbors should be estimating w_k^a.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Attack Model



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Attack Model

* How to get access to the neighbors' model $w_{k,i}$?

$$\boldsymbol{w}_{k,i-1} = \frac{\boldsymbol{\psi}_{k,i} - \mu_k \boldsymbol{u}_{k,i}^* \boldsymbol{d}_k(i)}{1 - \mu_k \boldsymbol{u}_{k,i}^* \boldsymbol{u}_{k,i}}$$

* Therefore, to deduce $w_{k,i-1}$, the attacker needs the knowledge of μ_k and streaming data $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$
  * μ_k can be obtained if it is uniform for all agents
  * $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$ are transferred from data fusion to agents – can be intercepted by the attacker
  * $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$ are sensed by agents – can be obtained by the attacker if it can get access to the sensor of the agents

Attack model

$$\boldsymbol{\psi}_{a,k,i} = \boldsymbol{w}_{k,i-1} + r_k^a(w_k^a - \boldsymbol{w}_{k,i-1})$$

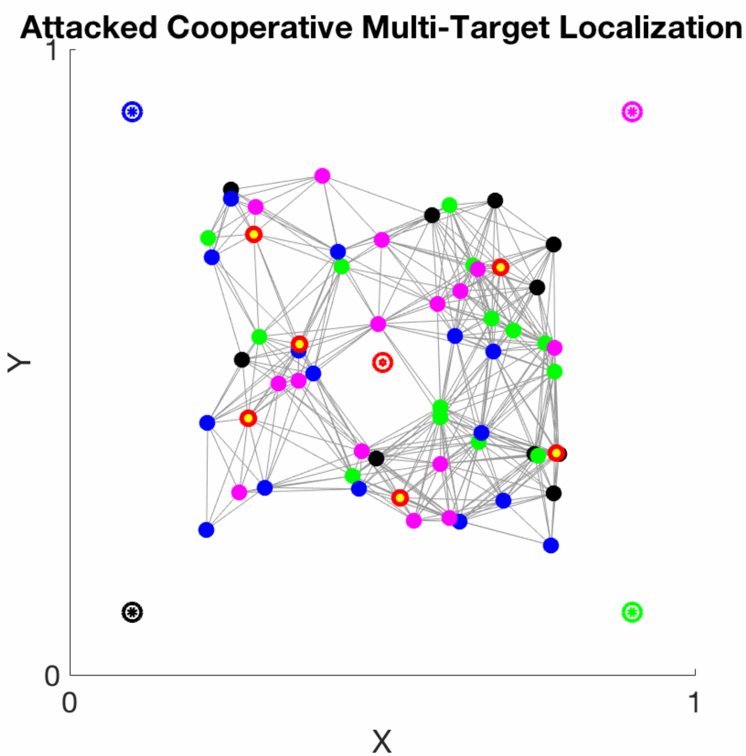
r_k^a should be small enough to satisfy $r_k^a(w_k^a - \boldsymbol{w}_{k,i-1}) \ll \mu_k \boldsymbol{u}_{k,i} e_k(i)$.

TABLE I: Streaming data representation in different problems

problem	w_k^0	$\boldsymbol{d}_k(i)$	$\boldsymbol{u}_{k,i}$
general distributed estimation / unsupervised clustering	estimated parameter	measurement received from data set	regression data received from data set
spectrum sensing / speech enhancement	signal transmitted by the source	signal received by agent k from the source	frequency-dependent attenuation factors over L frequency sample
target localization / biological design	target location	$\boldsymbol{d}_k(i) \triangleq \boldsymbol{\rho}_k(i) + \boldsymbol{u}_{k,i} z_{k,i}$, $\boldsymbol{\rho}_k(i)$ - sensed (noisy) distance measurement between agent k and target, $z_{k,i}$ - agent k's location	sensed (noisy) unit-norm direction vector pointing from the agent toward the target

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Select minimum set of agents to attack



Attackers
Attackers' selected target
Compromised agents (have converged to w_k^a)

- Attack model

$$\boxed{\psi_{a,k,i} = \boldsymbol{w}_{k,i-1} + r_k^a(w_k^a - \boldsymbol{w}_{k,i-1})}$$

- Objective
  - Attacker aims at compromising the entire network
- Select minimum set of agents to attack first
  - Find minimum dominating set of the graph
  - NP-complete and no efficient straight-forward solution!
- Attacker's way to approximate the solution

Algorithm Greedy Algorithm

1: $S := \emptyset$;
2: **while** $\exists$ white nodes **do**
3: choose $v \in \{x \mid w(x) = \max_{u \in V}\{w(u)\}\}$;
4: $S := S \cup \{v\}$;
5: **end while**

F. Kuhn and R. Wattenhofer. **Constant-Time Distributed Dominating Set Approximation**. In Proc. of the 22 nd Annual ACM Symp. on Principles of Distributed Computing (PODC), pages 25–32, 2003.

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Attack Detection

Agents will not be compromised if it processes data without cooperation

↓ strategy

Process data by two means: 1. without cooperation 2. proposed diffusion strategy.
When the estimation by diffusion strategy enters steady state, check the estimation difference between the two means.
Initialization: alarm(k) = 0 for all k.
If the difference exceeds a certain threshold, alarm(k) = 1.
If alarm(k) = 1, agent k will not trust in the estimation by diffusion strategy.

↓ use $\|\boldsymbol{w}_{\text{diff},k,i} - \boldsymbol{w}_{\text{diff},k,i-1}\| < \Delta$ to approximate steady-state

Can be satisfied before steady-state because of noise fluctuation

↓

False alarms

↓ remove false alarms

If alarm(k) = 1 and the difference is within a certain threshold, alarm(k) = 0.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Attack Detection

Algorithm 2 Resilient ATC diffusion strategy with adaptive combination weights

---

Set $\text{alarm}(k) = 0$, $\gamma_{lk}^2(-1) = 0$ for all $k = 1, 2, ..., N$ and $l \in N_k$

for all $k = 1, 2, ..., N, i \geq 0$ **do**

$\quad \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{\text{diff},k,i-1} + \mu_k \boldsymbol{u}_{k,i}^* (\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i} \boldsymbol{w}_{\text{diff},k,i-1})$

$\quad \gamma_{lk}^2(i) = (1 - \nu_k)\gamma_{lk}^2(i-1) + \nu_k \|\boldsymbol{\psi}_{l,i} - \boldsymbol{w}_{\text{diff},k,i-1}\|^2, l \in N_k$

$\quad a_{lk}(i) = \frac{\gamma_{lk}^{-2}(i)}{\sum_{m \in N_k} \gamma_{mk}^{-2}(i)}, l \in N_k$

$\quad \boldsymbol{w}_{\text{diff},k,i} = \sum_{l \in N_k} a_{lk}(i) \boldsymbol{\psi}_{l,i}$

$\quad \boldsymbol{w}_{\text{ncop},k,i} = \boldsymbol{w}_{\text{ncop},k,i-1} + \mu_k \boldsymbol{u}_{k,i}^* (\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i} \boldsymbol{w}_{\text{ncop},k,i-1})$

$\quad / * * * * * * * * * * \text{detection section} * * * * * * * * * * * /$

$\quad$ **if** $\|\boldsymbol{w}_{\text{diff},k,i} - \boldsymbol{w}_{\text{diff},k,i-1}\| < \Delta$

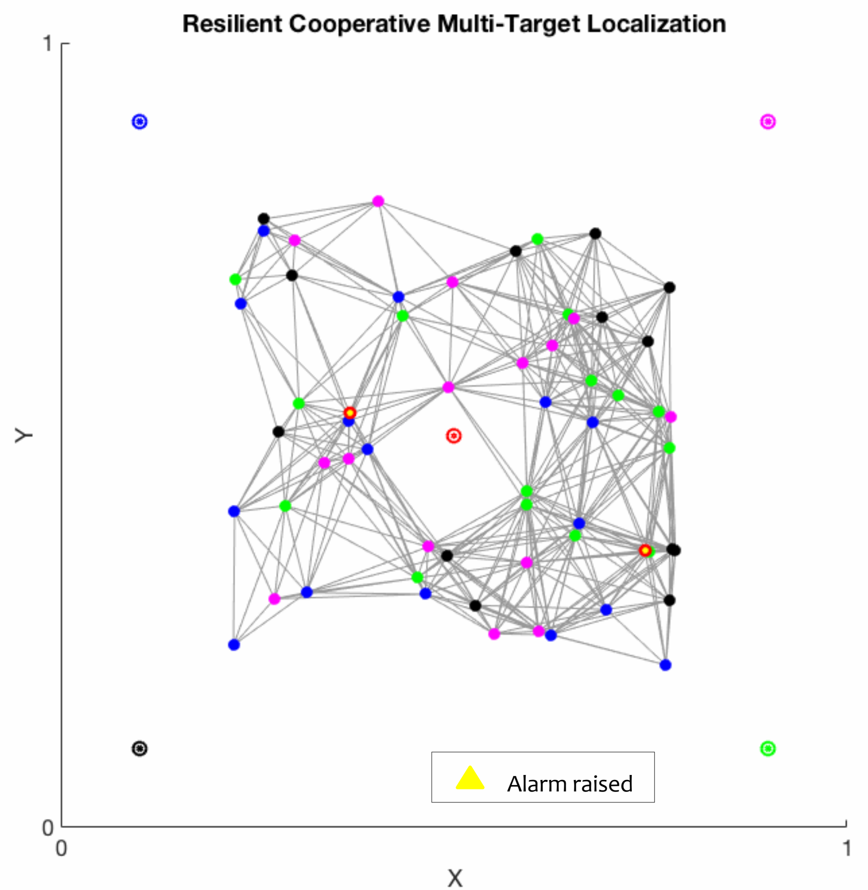
$\quad\quad$ **if** $\text{alarm}(k) = 0$ **and** $\|\boldsymbol{w}_{\text{ncop},k,i} - \boldsymbol{w}_{\text{diff},k,i}\| > \lambda$

$\quad\quad\quad \text{alarm}(k) = 1$

$\quad\quad$ **elseif** $\text{alarm}(k) = 1$ **and** $\|\boldsymbol{w}_{\text{ncop},k,i} - \boldsymbol{w}_{\text{diff},k,i}\| < \lambda$

$\quad\quad\quad \text{alarm}(k) = 0$

end for



FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Future work

* Time-varying distributed estimation case
* Large noise variance case

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Thank you!