Trust and Security Challenges for Networked Distributed Cyber-Physical Agent Systems On the edge

Carolyn Talcott
SRI International, USA

SBRC/WRNP PANEL
May 20, 2015

Context

Computing devices and Cyber-Physical Systems are everywhere

- Exciting opportunities for new infrastructure, medical, transportation, entertainment and other amazing apps
- Scary responsibilites to make systems that are usable, safe, responsive, trustworthy, …

Open interactive systems sense and affect their environments

- Must deal with uncertainty, faults, and partial knowledge
- Must adapt to resource constraints and disruptions in communication
- Require both autonomy and coordination

Vision – looking forward!

- Oceans of diverse, simple physically meaningful entities
  - declaratively defined behavior based on local information
  - collaborate opportunistically, locally and at a distance, to achieve diverse goals
- Redundancy for robustness and resilience

? How to design/build entities that realize the vision ?

Desiderata for CP agents

- **Localness**
  - agents must operate based on local knowledge
    - what they can observe / infer
    - what they can learn by knowledge sharing
- **Question:** How accurate / comprehensive does the local knowledge need to be in order to be able to (sufficiently) satisfy a given goal?

- **Safety/Liveness**
  - an agent should remain safe (healthy)
  - an agent should be able to act based on current information
    - should not require or need to rely on consensus formation
    - should be able to respond to change/threats in a timely manner
- **Question:** What does an agent need to monitor? How often?

- **Softness – for robustness and adaptability**
  - binary satisfaction is unrealistic
  - rigid constraints are likely to fail
  - consider a space of (feasible) solutions, rank them, and pick the best, which may differ according to situation
- **Question:** what time/space scope should be considered to (sufficiently) satisfy global goals by local actions?

Plan

POKS: Partially-Ordered Knowledge Sharing

- Infrastructure less, opportunistic communication model

A formal model of cyber physical agents

- Making the environment/physics explicit

Some example case studies

Challenges for open networked CPS

POKS: Partially Ordered Knowledge Sharing

Loosely-coupled Interaction through Sharing of Knowledge with a Partial Order (POKS)

- Knowledge items can be sensor readings, requests to actuators, locally computed solutions, community goals, etc.
- Knowledge items may be time-stamped
- Knowledge is shared opportunistically
  - provides delay/disruption-tolerant knowledge dissemination
  - does not require global coordination or infrastructure
  - supports entire spectrum between autonomy and cooperation

- Partial order captures
  - replacement – eliminate stale information, redundant goals
  - subsumption – logically redundant
- Primitives for knowledge sharing shield applications from the complexities of dealing with dynamic topologies, delays/disruptions, and failures

A model that makes the physical explict

We used the Maude rewriting logic system to define a formal framework that can be used to specify and analyze different behaviors of cyber-physical agents based on the POKS communication model.

The framework provides rules for communication (posting, propagating, and receiving knowledge) and templates for specifying agent actions using soft constraints and event handlers.

A system state consists of a set of agents of the form

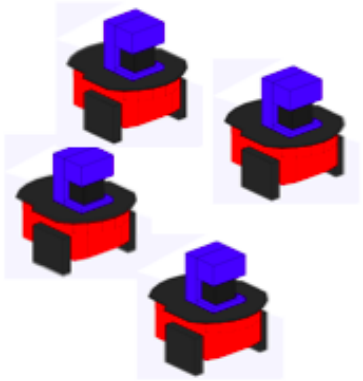
[id : class | envkb | localkb | cachedkb | events]

- localkb is the agents local knowledge
- cachedkb is knowledge to be opportunistically shared
- envkb represents the agents local physical environment
- events is the set of pending actions, tasks, knowledge to process

An agent system can be directly analyzed by rewriting using builtin or user defined strategies, or by searching for states having specified (desired or undesired) properties.

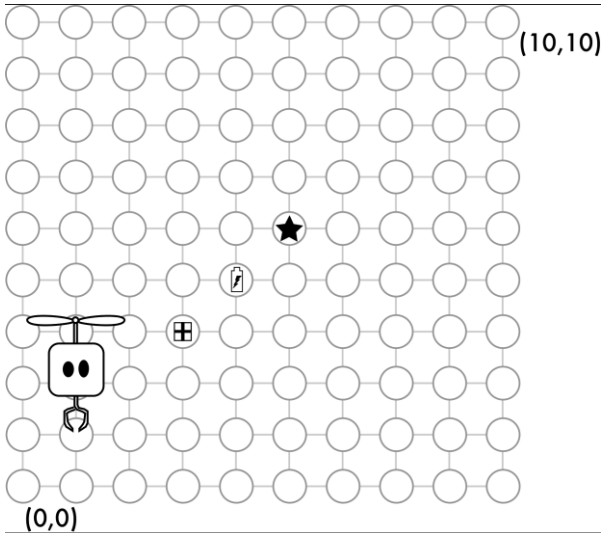
2D drone packet delivery system

- Situation: A city with packets to be picked up and delivered, and one or more courier drones. Drones get credit for successful delivery. Traveling uses energy. There are one or more charging stations and drones must not run out of energy.



- Packets to deliver next are ranked by energy cost subject to not running out of energy.

- Formal model specified starting with the Maude soft agent framework.



- With 2 drones and two packets, execution or search leads to a state with both packets delivered, each one by a different drone.

- The soft constraint system is `safe': If a bot starts in a safe state, it will never run out of energy.

3D Surveillance drone

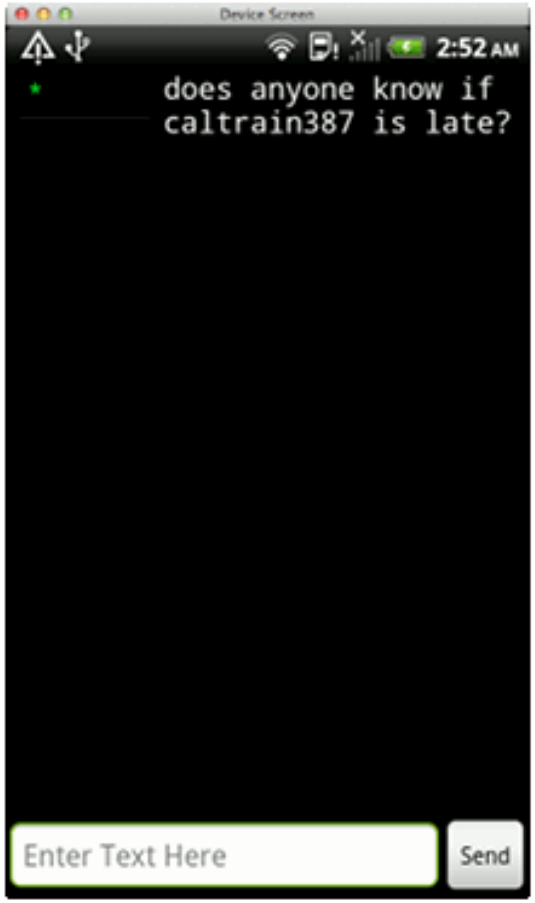
- There are one or more objects/areas to be photographed.
- 3D drones can move in one of 6 directions (U,D,N,S,E,W), and can take/store pictures.
- Clicking from higher elevation captures more targets but with less resolution.
- The objective is to gather pictures of all objects with the maximal resolution, given limited energy.
- Priorities
  - Stay safe -- always have enough energy to return home.
  - Take picture, if this increases quality of information
  - Move towards location with maximal opportunity
- Formal modeled specified starting with the soft agent framework.
- Drone actions are determined by soft constraints reflecting the priorities
- Scenario with 3 objects,
- After drone actions [click,U,click,D,N,N,click] the drone has pictures of all three objects.

CfChat: A Local Anonymous Chat

- An emphasis on (weakly) anonymous and local chatting, e.g., imagine standing at the train station platform and wondering if a train has been cancelled; being at an exposition sharing impressions; or in a emergency developing situation awareness.

- Cfchat allows the user to anonymously post a question and receive responses with an expectation that they originate from spatially close neighbors.

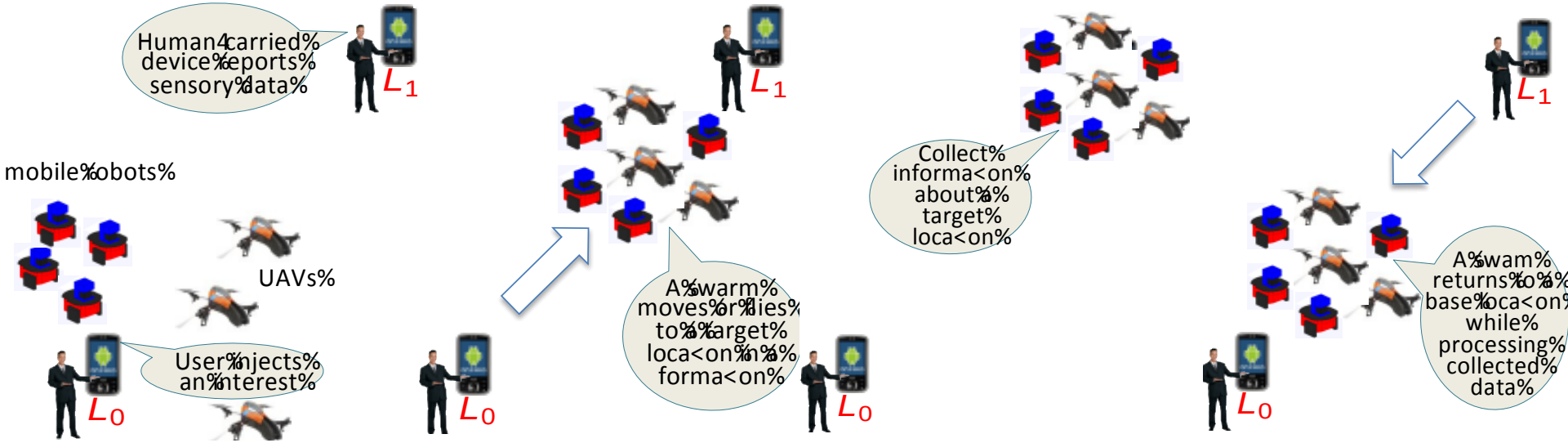
- Tagging messages allows the user to filter chat messages and tells the application which ones are important or should be disseminated.



Self-organizing Cyber-Physical Ensembles

A case study of an adaptive wireless network

- A swarm of programmable ground robots and UAVs that perform a distributed surveillance mission e.g., to achieve situation awareness during an emergency
- By moving or flying to a suspicious location, collecting information, and returning to a base location in a formation that creates an effective sensing grid.
- Involves human-carried computing/communication devices such as smart phones that collect/report sensor data and inject users' interests into the system.



What is missing

The POKS communication model makes developing applications involving dynamic, opportunistic interactions, without relying on infrastructure, simple in principle.

Within the model we can study questions such as

- Under what conditions are the local solutions good enough?
- Under what conditions would it not be possible to satisfy given goals?
- What quality of knowledge is needed for satisfactory solution/behavior?
- What frequency of decision making is needed so that local solutions are safe and effective?

BUT

- The POKS model is intentionally weak, making minimal assumptions about the environment.
- To make progress at the design level languages/models assume agents are honest, and knowledge is propagated correctly. We explicitly exclude considering insincere or malicious agents in our current formulation.

Removing the assumptions

Primitives are needed to deal with real world communication.

It would be interesting to use a CfChat-like app in emergency or disaster situations.

Responders could use CfChat to develop a useful if not perfect model of the situation.

For this to work mechanisms to establish (sufficient) trust without relying on infrastructure are needed.

- Verifiable / trustable assertions about capabilities, available resources, situations
- Crowd sourcing, but how to bootstrap?
- Anonymous identity? Key for CfChat is anonymity
- Trust is specific (trust for X)

Collaborative games are interesting for Museums to attract visitors and make visiting both fun and educational. For example, teams compete to `collect' artifacts and solve puzzles

POKS guarantees are needed:

- Communication abstractions – local transactions, soft synchronization – for game initiation.
- Unique copies of digital artifacts

Summary

- Potential applications of newly emerging technologies
  - Pico-satellite constellations
  - Networked balloons, buoys
  - CPS educational games
  - Smart environments – gardens, malls, conference venues, cities, …
  - Opportunistic IOT

- Issues
  - Building trust
  - Sufficient security
  - Situation awareness, detection and diagnosis of problems

That’s all folks!
???