Developing Automated Intelligence Collection Plans from Probabilistic Behavior Estimates

Georgiy Levchuk
Scott Galster
Krishna Pattipati

Presented at 2009 ICCRTS

Date: 06/17/2009

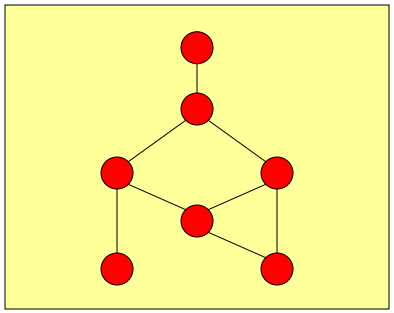
www.aptima.com
Boston ▪ DC ▪ Dayton

APTIMA HUMAN-CENTERED ENGINEERING

Agenda

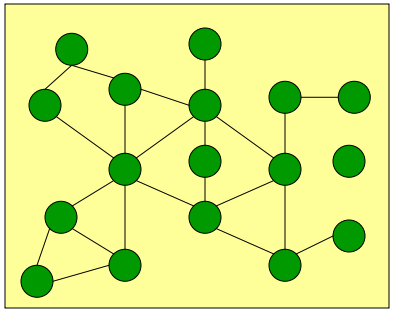
- Definitions
- Problem
- Approach
- Results

Definitions-1



- **RED**
  - adversaries, target of analysis

- **BLUE**
  - friendly forces, users of the tool, analysts



- **GREEN**
  - "normal" (local) population, not RED/BLUE

- **Resources**
  - people, materials, physical infrastructure, information, etc.

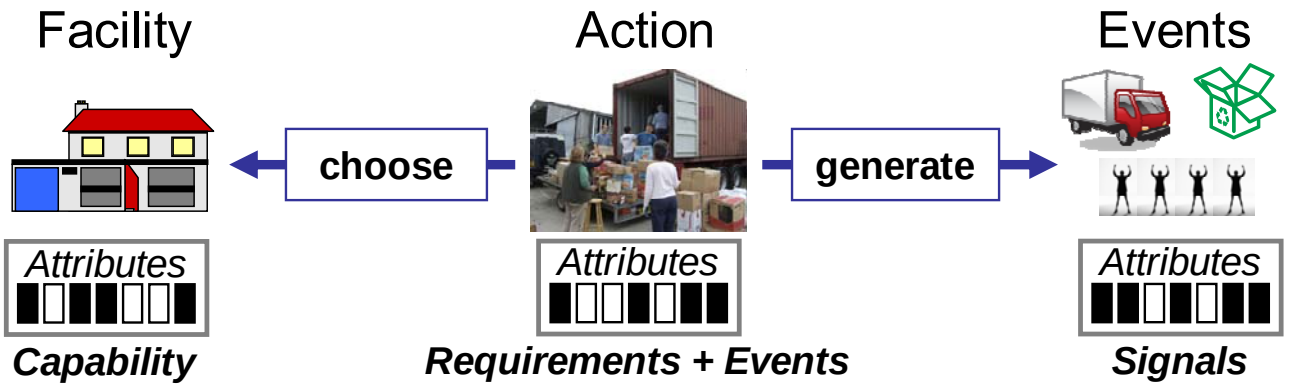
Definitions-2

- **Actors**
  - people, moving objects (e.g., cars), places
- **Actions**
  - performed by actors
- **Attributes**
  - quantitative description for actors (capabilities, preferences, objectives) and actions (requirements, outcomes)

Definitions-3:
Attributes & specification of the relations between observations, actors, and actions

Examples of attributes:

- **Choice/req-s attributes:** why would a facility be used to carry an activity
  - Example: "*assemble weapons in building with electricity supply and extra generator*"
- **Signal/event attributes:** what data might be observable if the activity is taking place
  - Example: "*weapons assembly activity would generate a spike in electricity use, which might be observed if electricity flow is monitored*"



Data: Choice Attributes | **Model: Choice & Signal Attributes** | **Data: Signal Attributes**

Definitions-4

- **Sensors / data sources**
  - HUMINT, SIGINT, IMINT, MASINT, OSINT, GeoINT
- **Observations**
  - quantitative and qualitative data obtained by sensors about actors and actions
- **Behaviors**
  - (patterns of) actions, either oriented by objective or not

Definitions-5: Behavior Types

- Single objects…

entering building

digging a hole

- Multiple objects…

meeting

playing

- Static objects…

gas station

kindergarten

Definitions-6

- **Networks**
  - actors, their roles, and their relationships
- **Missions / scenarios**
  - plans composed of patterns of actions oriented by an objective
- **Behavior Signature**
  - network(s) + mission(s)

Definitions-7: Mission = Action Precedence Graph

Model network nodes

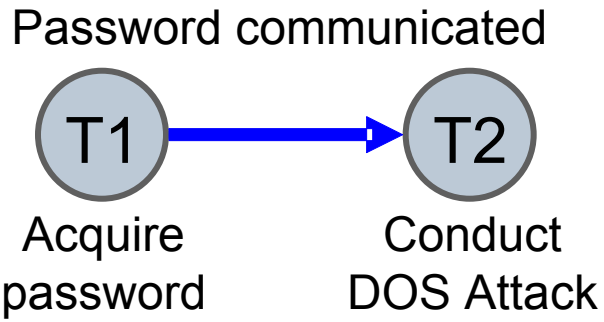
- Actions/tasks to be performed by actors

Model network links

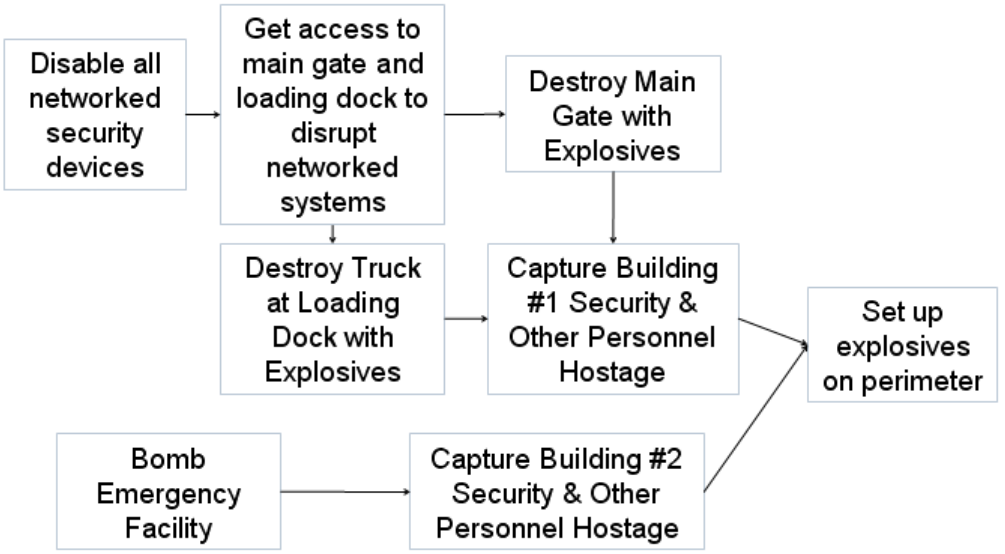
- Precedence, info, material flow

Attributes

- Requirements for task/activity resources
- Capabilities of actors/facilities needed for carrying the tasks
- Utilities & preferences

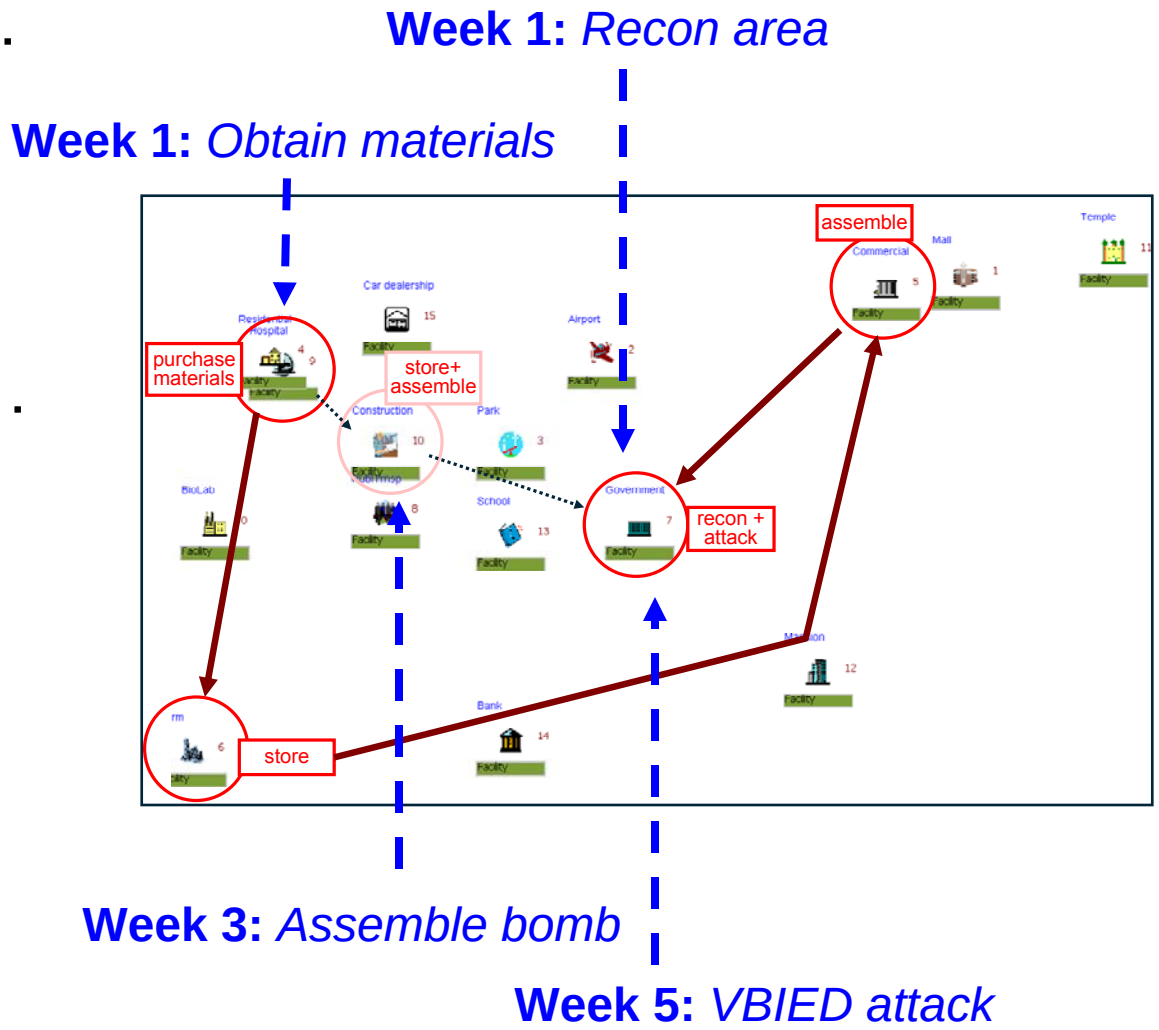


Example of RED Mission



Definitions-8: Missions = Coordinated Behaviors

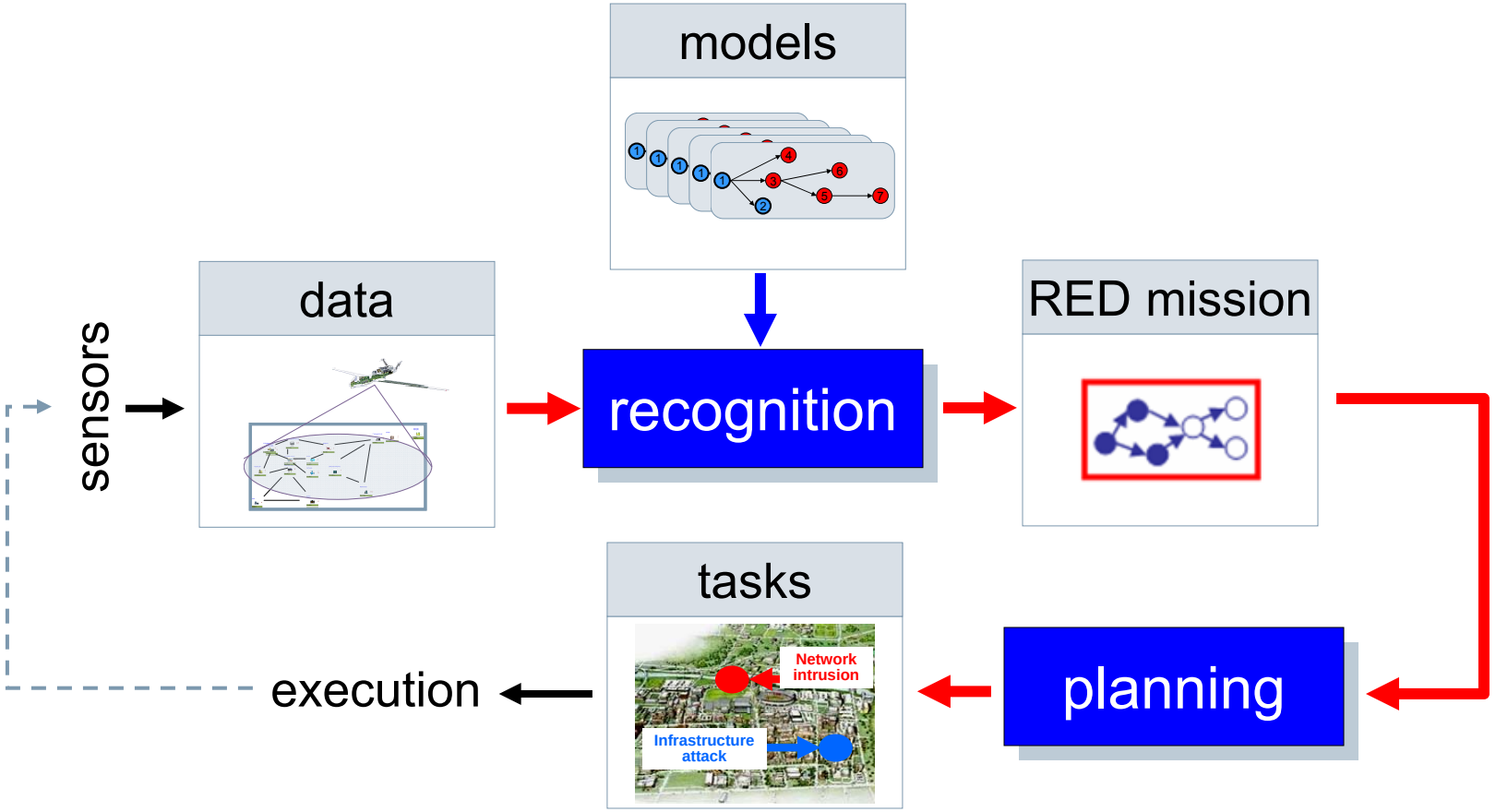
- Multiple places…
- Different actors…
- Different times…



Definitions-9

- **Data**
  - all observations
- **Models**
  - known patterns of behavior, missions, and network (sub)structures

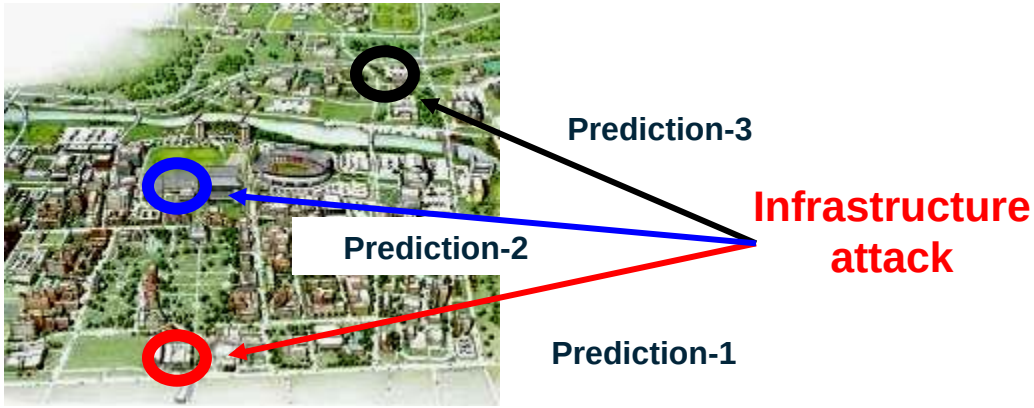
The Problem



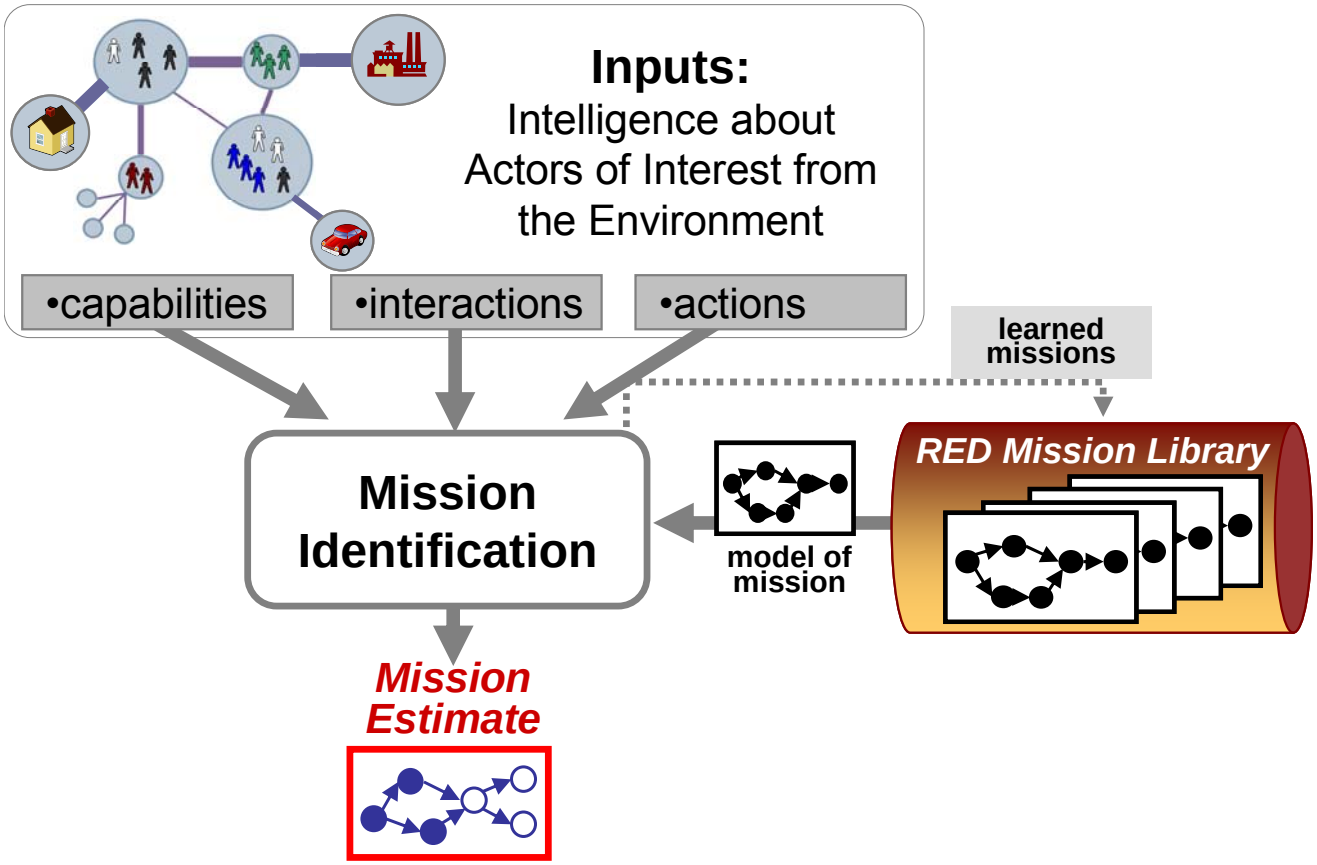
- Integrate collection planning with probabilistic situation assessment models

General Needs for Collection Planning

- Improve reliability/robustness of situation assessment
  - **disambiguate among current predictions**
- Identify critical missing information
- Prioritize collection actions to achieve highest information gain under cost constraints



Prediction as Hypotheses Testing

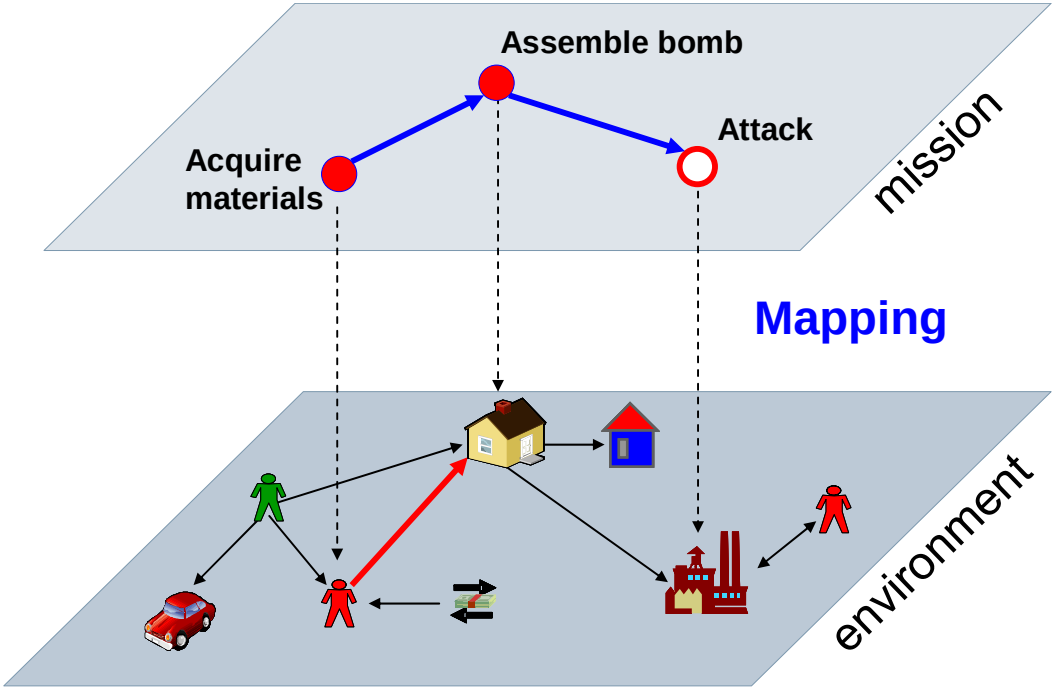


Addressed Need: Identification of Critical Information

- Prediction consists of
  - RED mission
  - State of RED mission
  - Mapping of RED mission to areas and actors
  - Probability of mission & mapping

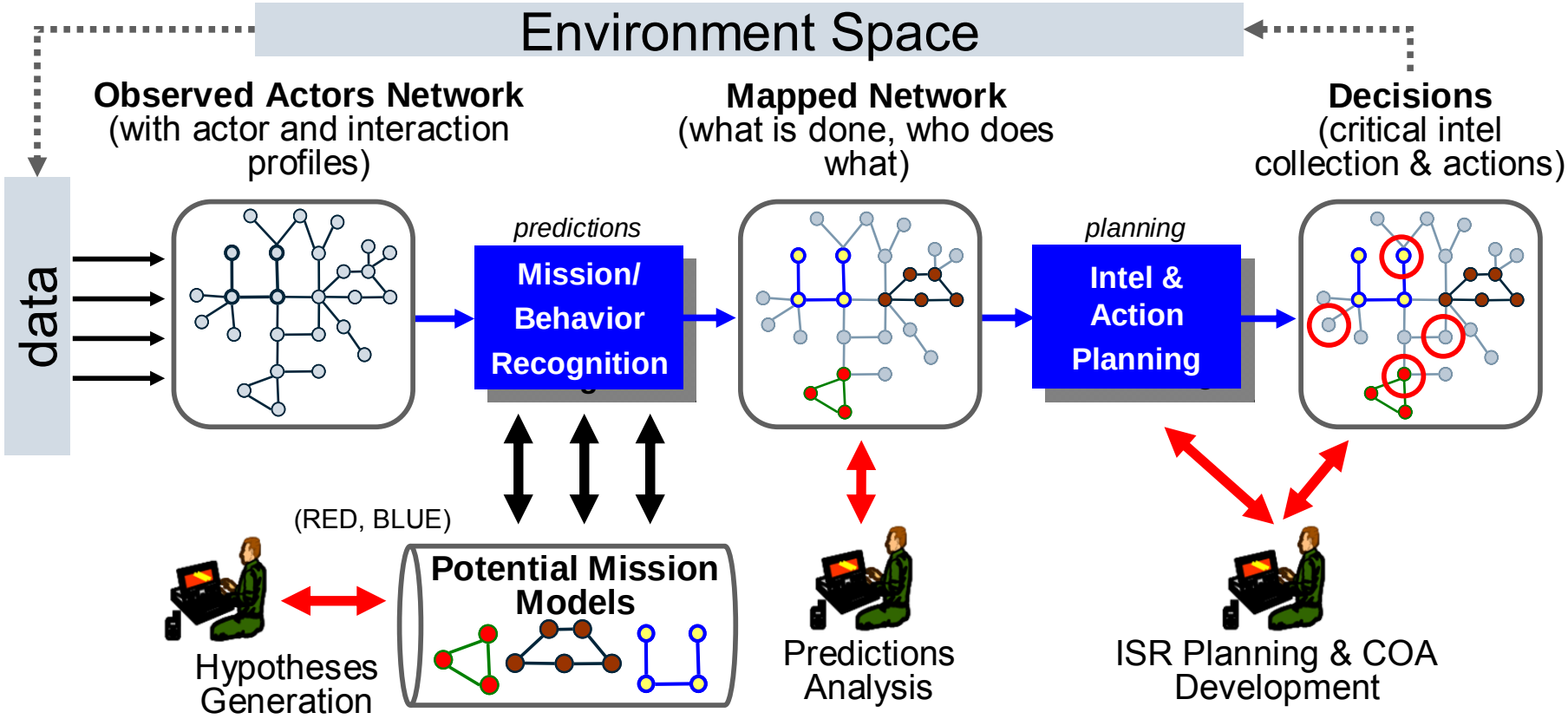
- Need to disambiguate
  - Different RED missions
  - Different RED mission states
  - Different RED mission mappings

- Prediction defines the task mapping for each actor



APTIMA HUMAN-CENTERED ENGINEERING

Our Approach Workflow



The Meaning of "Probabilistic Disambiguation"

- Simple disambiguation
  - Suppose have multiple information elements (aka predictions) that are defined via vectors of features $\bar{v}^i = (x_1^i, \ldots, x_n^i), i = 1, \ldots, M$
  - Then if a feature ***k*** is
    - the same for all elements, i.e. $x_k^i = x_k^j, i \neq j$, then it is NOT disambiguating
    - different for all elements, i.e. $x_k^i \neq x_k^j, i \neq j$, then it is most disambiguating
  - Example:

$$\begin{matrix} \bar{v}^1: \\ \bar{v}^2: \\ \bar{v}^3: \\ \bar{v}^4: \end{matrix} \begin{bmatrix} 0 & 2 & 1 & 0 \\ 5 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

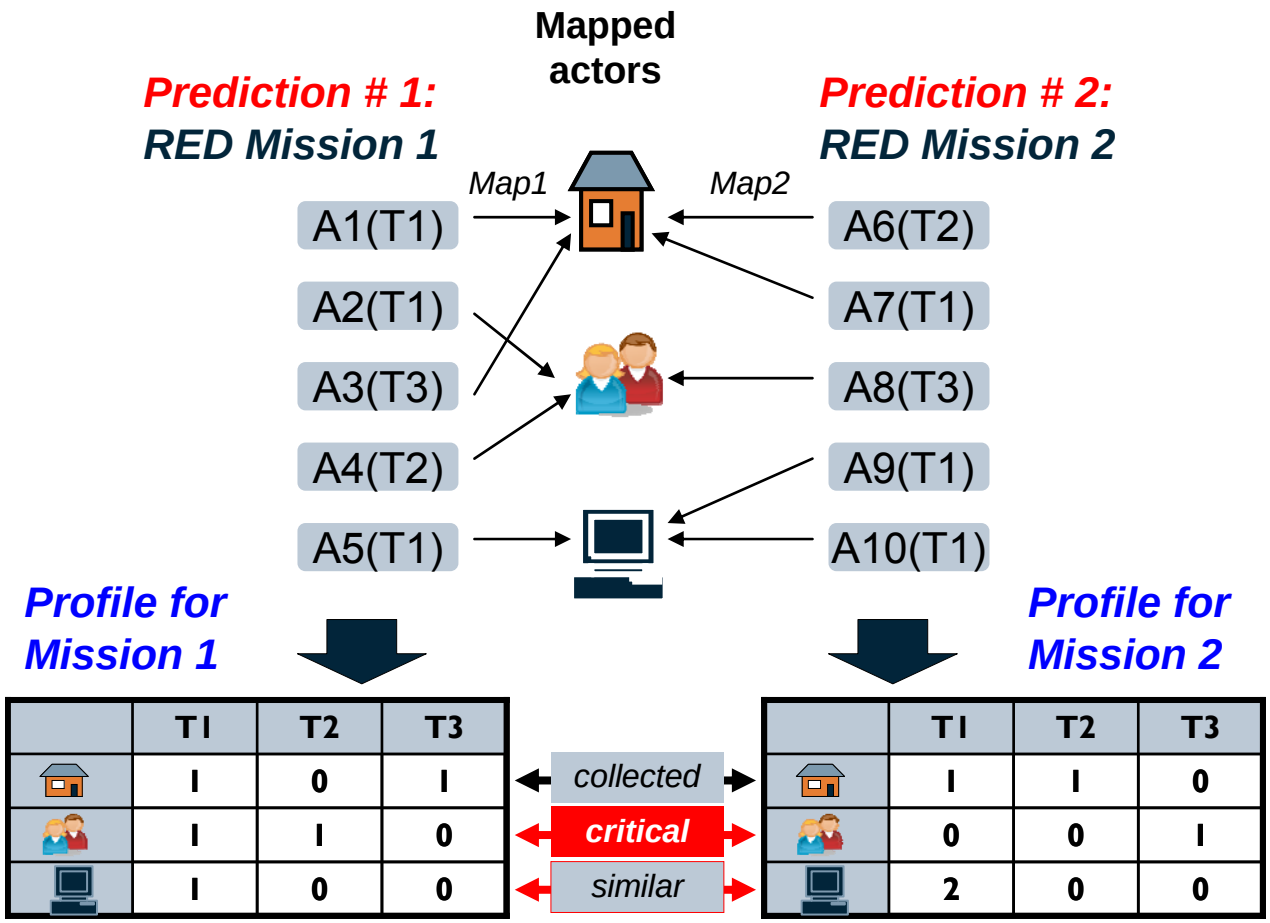
MAX disambiguation

NO disambiguation

- Probabilistic disambiguation
  - Weights on the "benefit" of disambiguating certain elements

Idea for Plan Design: Prediction's Behavior Signature Profiles

- Each actor/area is target for intel collection
- For each prediction, we develop actor/area profiles based on mapped task features
- This allows us to see differences that a collection at the actor can make (how many predictions have distinct profile at the actor)



Formal Plan Design Process

- Generalizations to information-theoretic planning
  - Actors = information elements
  - Action types = features
  - Action mapping = actor feature vectors
- Objective:
  - Maximize **Information Gain** (minimize entropy) of collection actions

$$gain(O) = \underbrace{H(G_M, S_M \mid G_D)}_{\text{current information}} - \underbrace{H(G_M, S_M \mid G_D, O)}_{\text{new information}}$$

Compute using probabilities and prediction disambiguation counts

- Process:
  - Prioritize information elements in the order of increased information gain (reduced entropy) constrained by the cost of commensurate collection actions
  - Cluster related collection actions
  - Generate the plan as a decision tree with each decision nodes defined with information collection action and each outgoing link associated with possible outcome of collection

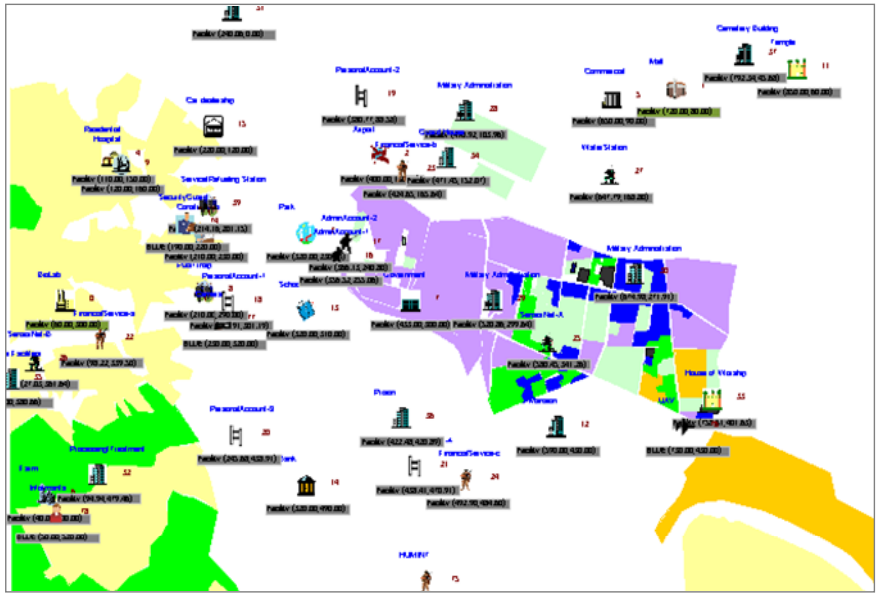
APTIMA HUMAN-CENTERED ENGINEERING

Summary of Accomplishments

- Used several **real-world data sets** supplemented by synthetic data with ground truth for evaluating the technology
- Showed that ISR collection planning **improves the accuracy of situation assessment** by targeting the information collection most critical to current predictions

Example Data Set: Terrain

- Terrain included buildings and actors of various types
- Information (possibly noisy) about their capabilities / objectives was available



(a) Area Layout for Dataset

Area	Function
BioLab	Plant
Mall	Infrastructure
Airport	Infrastructure
Park	Social
Farm	Infrastructure
Government	Government
FinancialService	Infrastructure
Oil/Gas Facilities	Military
SensorNet	NetworkNode
Military Administration	Military
WaterStation	Infrastructure
AdminAccount	Government

(b) Example of Building List and Functions

Example Data Set: Actors and Actions/Tasks

- Variety of actions and actors was modeled in the dataset

	Resource Requirements													*Target Requirements*											
Role	SZ	SEC	STR	MAT	TEC	KNW	MON	REC	POIS	AINF	PINF	BACT	CSENS	SZ	SEC	STR	MAT	TEC	KNW	MON	REC	POIS	AINF	PINF	BACT	CSENS
Acquiring poison	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Recon	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Storing explosive materials	0	0	0	1	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
Assemble bomb	0	0	0	1	0	1	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0
Insert Trojans to Capture Additional Passwords and Changes	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Siphon Funds from Compromised Accounts and Change Passwords to Lock out Users and Admins	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Gain control over network to disable/manipulate sensors/monitoring capabilities/system	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Create false threat of bomb attack against government building	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(a) RED Model Tasks/Activities

	Capabilities												
Facility	SZ	SEC	STR	MAT	TEC	KNW	MON	REC	POIS	AINF	PINF	BACT	CSENS
BioLab	2	0	1	0	1	0	0	0	1	0	0	0	0
Mall	4	1	3	0	0	0	2	0	0	0	0	1	0
Airport	10	0	3	3	2	0	0	0	0	0	0	0	1
Park	3	3	0	0	0	0	0	0	0	0	0	0	0
Residential	1	5	1	0	0	0	0	0	0	0	0	0	0
Commercial	3	2	3	1	2	0	1	0	0	0	0	0	0
AdminAccount	0	0	0	0	0	0	0	0	0	1	0	0	0
SensorNet	0	0	0	0	0	0	0	0	0	1	0	0	1

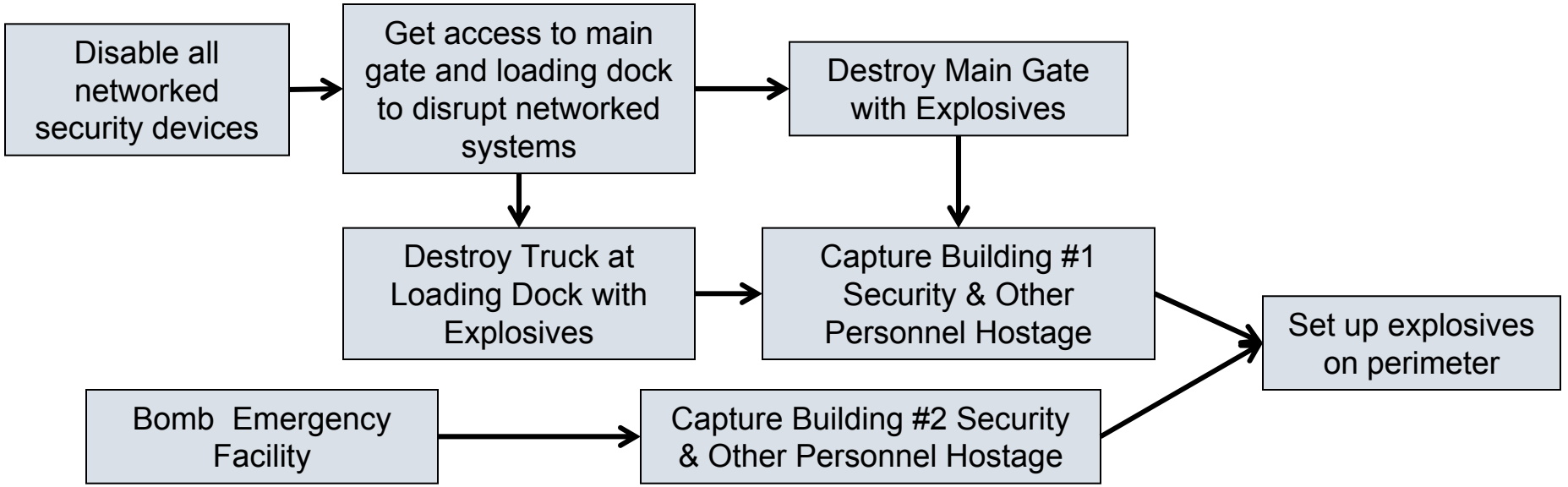
(b) RED Areas/Facilities

	Capabilities												
Role	SZ	SEC	STR	MAT	TEC	KNW	MON	REC	POIS	AINF	PINF	BACT	CSENS
SecurityDetail	0	1	0	0	0	0	0	0	0	0	0	0	0
Hackers	0	0	0	0	0	0	0	0	0	1	1	1	0
Attacker	0	0	0	0	0	0	0	0	1	0	0	0	0
Financier	0	0	0	0	0	0	1	0	0	0	0	0	0
Recon	0	0	0	0	0	0	0	1	0	0	0	0	0
Bombmaker	0	0	0	0	0	1	0	0	0	0	0	0	0

(c) RED Actors

Example Data Set: RED Mission Hypotheses

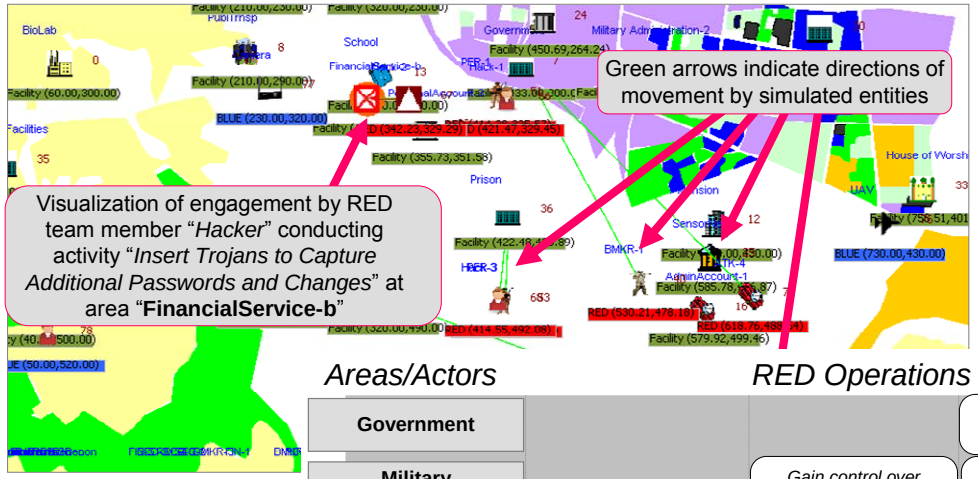
- Several hypothetical RED missions were designed for dataset, for example:



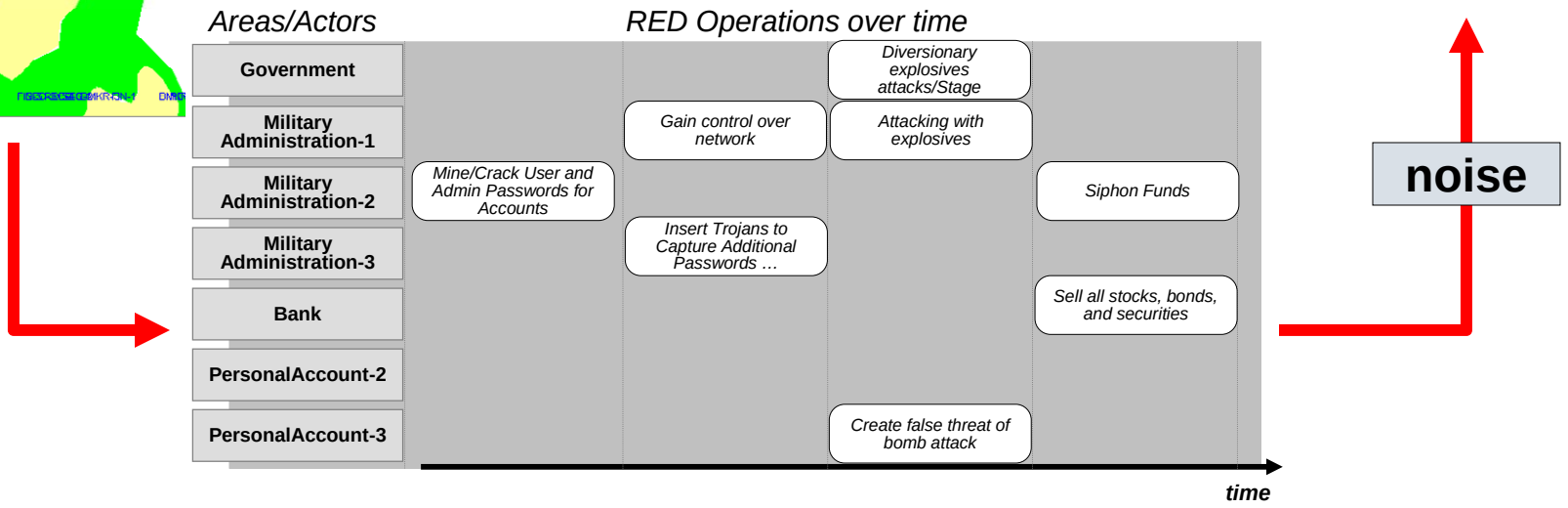
Mission: Airport Capture/Hostages

Example Data Set: Observations

- Simulated events have been converted into actor profiles thru noise component



	Capability of Area/Actor						Current events of Area/Actor								
Actors/Areas	VAL	GOV	AINF	PINF	CSVC	CINFR	KNW	ATK	AINF	PINF	BACT	CSVC	CINFR	HACK	PER
Airport	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Park	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Residential	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Commercial	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Government	3	1	1	1	1	1	1	0	0	0	0	0	0	0	1
Mansion	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
School	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Bank	1	0	1	1	1	0	0	0	0	1	0	1	1	1	0
AdminAccount-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PersonalAccount-1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
PersonalAccount-2	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0
PersonalAccount-3	0	1	1	0	0	0	0	0	1	0	0	0	1	0	0
PersonalAccount-4	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
FinancialService-a	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
SensorNet-A	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
Military Administration-1	5	0	0	1	0	1	0	1	1	0	0	0	0	1	1
Military Administration-2	5	0	0	1	0	1	0	0	0	1	1	0	1	1	0
Military Administration-3	5	0	0	1	0	1	0	0	1	0	0	0	0	1	0
Telecommunications	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Processing/Treatment	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0



Example Data Set: Noise Process

- Several types of errors introduced into observations

(a) True Attribute Vector

	Capabilities									Current Events								
Facility	VAL	TRS	STR	REC	ATK	MON	POIS	SEC	TEC	VAL	TRS	STR	REC	ATK	MON	POIS	SEC	TEC
BioLab	0	0	0	0	0	0	1	0	1	0	1	1	0	0	1	0	0	0

Static pre-mission intel **Dynamics events/intel**

	Capabilities									Current Events								
Facility	VAL	TRS	STR	REC	ATK	MON	POIS	POIS	TEC	VAL	TRS	STR	REC	ATK	MON	POIS	SEC	TEC
BioLab	**1**	0	0	0	0	0	**2**	0	**0**	0	1	**0**	0	0	1	0	0	0

Irrelevant attribute

Attribute error

Attribute miss

Event miss

(b) Observed Attribute Vector

APTIMA HUMAN-CENTERED ENGINEERING

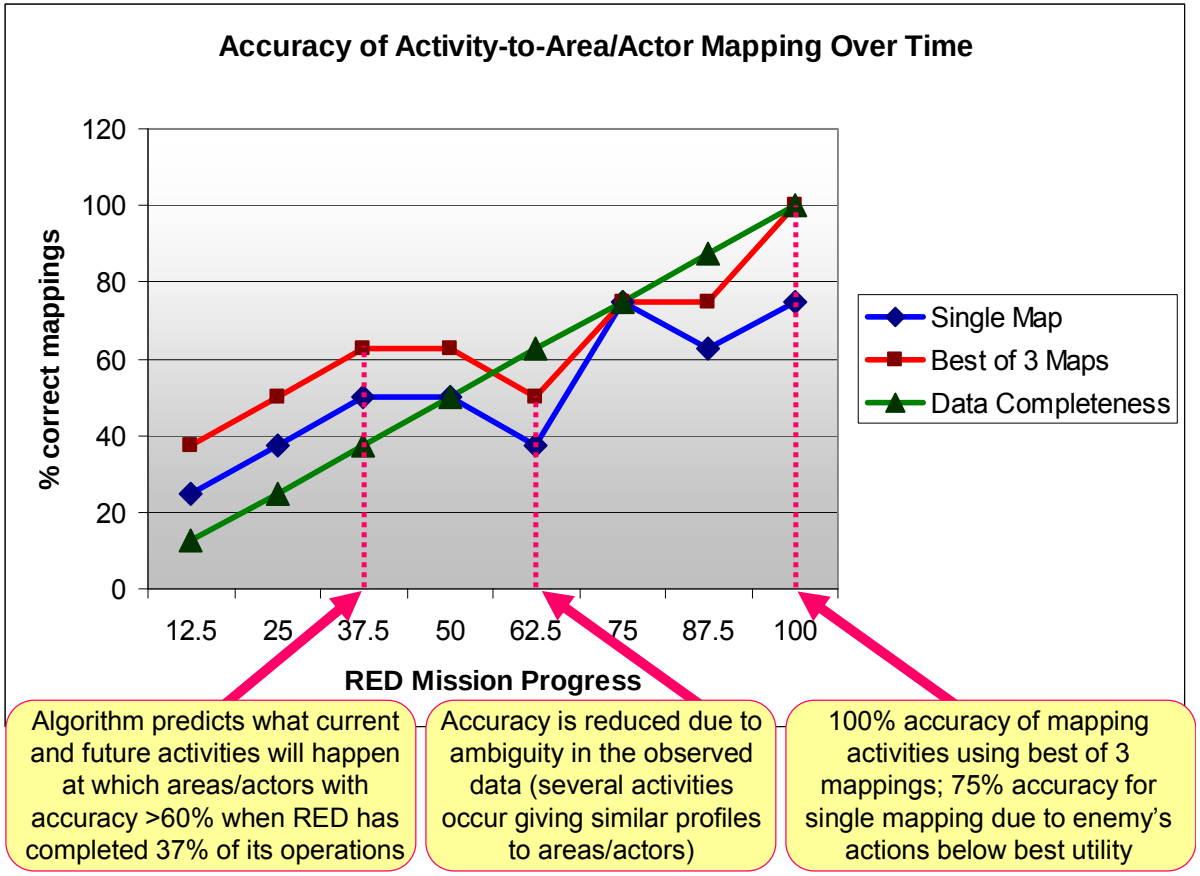
Example of Analysis: Predictions/mappings

Task Name	**Mapped Area/Actor (1)**	**Mapped Area/Actor (2)**	**Mapped Area/Actor (3)**
Attacking with explosives	**Military Administration-1**	**Military Administration-1**	**Military Administration-1**
Diversionary explosives attacks/Stage	**Government**	**Government**	**Government**
Mine/Crack User and Admin Passwords for Accounts	**PersonalAccount-2**	**PersonalAccount-3**	**Bank**
Insert Trojans to Capture Additional Passwords and Changes	**Military Administration-3**	**Military Administration-1**	**Bank**
Create false threat of bomb attack	**PersonalAccount-3**	**PersonalAccount-3**	**PersonalAccount-3**
Sell all stocks, bonds, and securities	**Bank**	**Bank**	**Bank**
Siphon Funds	**Military Administration-2**	**Military Administration-2**	**Military Administration-2**
Gain control over network	**Military Administration-1**	**Military Administration-1**	**Military Administration-1**
% correct	100%	75%	75%

Mapping of Actions to Actors (yellow cells indicate incorrect predictions)

Example of Analysis: Sensitivity of Predictions

- Accuracy goes down when receive more but ambiguous observations
- Indicates importance of collecting data that disambiguates rather than data that increases the confusion



Example of Analysis: Profiles & Critical Information

- Example behavior signature profiles for the analyzed dataset
  - **Task profiles** = mapped task types (high-level info element disambiguation analysis)
  - **Feature/event profiles** = aggregated task requirements (detailed disambiguation analysis)

	Task Profile	Feature/Event Profile								
Map1		KNW	ATK	AINF	PINF	BACT	CSVC	CINFR	HACK	PER
Military Administration-1	[1,0,0,0,0,0,0,1]	0	1	1	0	0	0	0	1	1
Military Administration-2	[0,0,0,0,0,0,1,0]	0	0	0	1	1	0	1	1	0
Government	[0,1,0,0,0,0,0,0]	1	0	0	0	0	0	0	0	1
PersonalAccount-2	[0,0,1,0,0,0,0,0]	0	0	0	0	0	0	0	1	0
PersonalAccount-3	[0,0,0,0,0,0,0,0]	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Bank	[0,0,0,0,0,1,0,0]	0	0	0	1	0	1	1	1	0
Map2										
Military Administration-1	[1,0,0,1,0,0,0,1]	0	1	2	0	0	0	0	2	1
Military Administration-2	[0,0,0,0,0,0,1,0]	0	0	0	1	1	0	1	1	0
Government	[0,1,0,0,0,0,0,0]	1	0	0	0	0	0	0	0	1
PersonalAccount-2	[0,0,0,0,0,0,0,0]	0	0	0	0	0	0	0	0	0
PersonalAccount-3	[0,0,1,0,1,0,0,0]	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**0**
Bank	[0,0,0,0,0,1,0,0]	0	0	0	1	0	1	1	1	0
Map3										
Military Administration-1	[1,0,0,0,0,0,0,1]	0	1	1	0	0	0	0	1	1
Military Administration-2	[0,0,0,0,0,0,1,0]	0	0	0	1	1	0	1	1	0
Government	[0,1,0,0,0,0,0,0]	1	0	0	0	0	0	0	0	1
PersonalAccount-2	[0,0,0,0,0,0,0,0]	0	0	0	0	0	0	0	0	0
PersonalAccount-3	[0,0,0,1,0,0,0,0]	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**0**
Bank	[0,0,0,0,0,1,0,0]	0	0	0	1	0	1	1	1	0

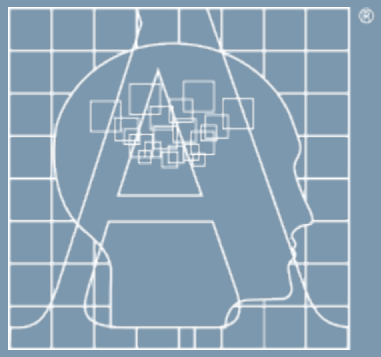
Feature/event profile for "**Military Administration-1**" looks the same for all three mappings – additional data collection will not disambiguate these mappings

"**PersonalAccount-3**" can disambiguate all three mappings. It has 0-feature vector for mapping 1, and its non-zero feature vectors for mappings 2 and 3 are distinguished by feature type/event "HACK"

"**PersonalAccount-2**" cannot disambiguate all three mappings as it has same 0-feature vectors for mapping 2 and 3

Conclusions

- Developed approaches for automating integration between adversarial reasoning / situation assessment and ISR collection planning technologies
- Obtained high accuracy of behavior/mission pattern recognition and activity mapping for large levels of data uncertainty
- ISR collection planning improves the accuracy of the situation assessment further by targeting the information collection most critical to current predictions
- We have illustrated the process of situation assessment and ISR planning on the example dataset



APTIMA®

HUMAN-CENTERED
ENGINEERING