Cyber Trust and Suspicion

Eunice E. Santos
Institute of Defense & Security
University of Texas
El Paso, TX
eesantos@utep.edu

Institutional Members

- (Lead:) Institute of Defense & Security, The University of Texas at El Paso
- Assured Information Security, Inc.
- Dartmouth College
- Laureate Institute for Brain Research
- Syracuse University
- The University of Texas Health Science Center at Houston
- The University of Tulsa
- (Partner) 711 HPW/RHC

Motivation

- Trust and suspicion are critical components in cyberspace operations (CO) especially in regards to the information technology systems that are involved in such operations regardless of whether they are defensive or offensive in nature.
- The human is at the center of CO being the primary entities susceptible to trust issues and suspicion.
- How the humans and human organizations react to trust and suspicion plays a significant role in shaping the outcome of CO on both sides of the mission.

Goals

- Provide the fundamental research in building the foundations for analyzing and understanding the impact of human trust and suspicion in the cyber war environment
- Explore a multi-pronged and multidisciplinary approach to address the myriad of factors and variations inherent in the human operator
- Develop computational, neural science and social science constructs in order to tease apart the complexities of this problem space

Objectives

- Developing a model of insider behavior that accounts for and explains the social, cultural, and emotional basis for trust and suspicion especially its impacts on insider threat.
- Research and identify biomarkers of cyber trust for the selection of targeted training and interface/alert interventions.
- Systematically demonstrate and examine how human performance affects cyber security operations with humans in the loop, and explore how such effects can be mitigated or exploited in order to achieve a higher-level of security.
- Conduct human subject studies (where subjects are equipped with non-invasive sensors) to provide real-time predictions about the changing level of trust and suspicion experienced by subjects while they conduct tasks that are designed specifically to test hypotheses stemming from the other team members' research.
- Assess, attribute, and manipulate operator suspicion through cyber means and demonstrating formal models of suspicion.

5 Principal Thrusts

- Thrusts serve as seeds to explore the different aspects of this space which can further enhance our understanding through eventual cross-fertilization of ideas:

1. A Social, Cultural, and Emotional Basis for Trust and Suspicion: Manipulating Insider Threat in Cyber Intelligence & Operations
2. Targeted Interventions Derived from Biomarkers of Cyber Trust
3. A Human-Centric Approach to Cyber Trust and Suspicion
4. Using Non-invasive Sensors to Predict Trust and Suspicion in Human Operators
5. Assessing, Attributing, and Manipulating Operator Suspicion

A Social, Cultural, and Emotional Basis for Trust and Suspicion: Manipulating Insider Threat in Cyber Intelligence & Operations

THRUST 1

Thrust 1 Team

- (Thrust PI:) Dr. Eunice E. Santos, Founding Director, Institute of Defense & Security, The University of Texas at El Paso
- Dr. Eugene Santos, Jr., Professor, Thayer School of Engineering, Dartmouth College
- Dr. John Korah, Research Assistant Professor, The University of Texas at El Paso

Goals

- By combining computational and social science constructs, our goal is to develop a model of insider behavior that accounts for and explains the social, cultural, and emotional basis for trust and suspicion especially its impacts on insider threat.

Target Questions

a) How can different people be swayed (or sway others) based on trust or suspicion?
b) How and why do individual socio-cultural characteristics, group size, information sharing paradigms and events affect operational cohesion?
c) Is it possible to detect significant drops in situational awareness, or when the level of trust is inappropriate in a given context?
d) What are the critical inter-relationships between information manipulations, emotional responses, situational awareness, influences on decision-making, and associated changes in task performance/cyberspace operations?
e) How do complex multi-scale and multi-level factors in cyberspace operations impact insider threat detection and manipulation?
f) Can we unify this into a single overarching framework of social, cultural, and emotional factors underlying trust and suspicion for manipulating insider threat?

Anticipated Results

- Methodology that developers and operators could use to better understand and exploit insider threat by providing the social, cultural, and emotional basis of insider behavior and the impacts of trust and suspicion on cyberspace operations
- Understand why they occur, how they occur, and how they can be mitigated, managed, or manipulated
- To date, there has been little or no work in providing any unified/comprehensive treatment of the impacts of social, cultural, economic, political, and emotional factors (to name a few) underlying trust and suspicion in insider threat especially in complex systems/organizations involving multi-level and multi-scale effects and dynamics

Objectives

- Develop a model that explains how different people can be swayed (or sway others) based on the amount they are trusted.
- Study and develop a model of how individual socio-cultural characteristics, group size, information sharing paradigms and events affect group cohesion.
- Develop an approach to detect significant drops in situational awareness, or when the level of trust is inappropriate in a given context.
- Understand the relationships between information manipulations, emotional responses, situational awareness, sensemaking, influences on decision-making, and associated changes in task performance/cyberspace operations.
- Explore and define the mechanisms of manipulating a cyberspace information environment and explain how it effects changes in task performance/cyberspace operations.
- Understand and account for complex multi-scale and multi-level factors in cyberspace operations as it impacts insider threat detection and manipulation.
- Define an overarching framework that unifies social, cultural, and emotional factors underlying trust and suspicion for manipulating insider threat.

The Insider

- From www.Miriam-Webster.com:
  - a person recognized or accepted as a member of a group, category, or organization: as (a) a person who is in a position of power or has access to confidential information, (b) one (as an officer or director) who is in a position to have special knowledge of the affairs of or to influence the decisions of a company

- Conducting cyberspace operations requires that your organization engage in *managing your insiders*, *mitigating your malicious insiders*, and *manipulating your opponent's insiders*.

It’s about the people!!

- Cyberspace operations and intelligence involve *people:*
  - People who have different motivations, goals, and intentions;
  - People who have different cyber abilities, tasks, and resources;
  - People who have different beliefs, culture, and politics;
  - People who form different groups, cliques, and organizations;
  - People who interact differently with friends, relatives, strangers, enemies, and organizations;
  - People who react differently to stress, chaos, and emotions; and, not the least,
  - People can change by themselves, are changed by others, and even changed by the cyber environment.

Approach Organization

- Insider intent modeling
- Multi-scale, multi-level, socio-cultural behavior modeling
- Emotion and decision-making (in collaboration with Dr. Michael Haas, 711th Human Performance Wing, AFRL)
- Measuring trust and suspicion in cyberspace operations (in collaboration with Dr. Leanne Hirshfield – Thrust 4)

Insider Intent Modeling

- Define a model to capture insider's **intent**
  - Focus on modeling the process of achieving a goal
  - Focus on capturing intent "on-the-fly"
  - Focus on evaluation of intent modeling based on synthesized data and human generated data

Insider Intent Model

- **Insider intent** = Goals + Actions + Commitment
  - Our intent model consists of 3 components that are designed to capture intent:
  - **Foci**: "**What** is the working space of the insider and what they are concentrating on?"
  - **Rationale**: "**Why** does the insider have these foci?"
  - **Action**: "**How** are the insider's goals accomplished?"

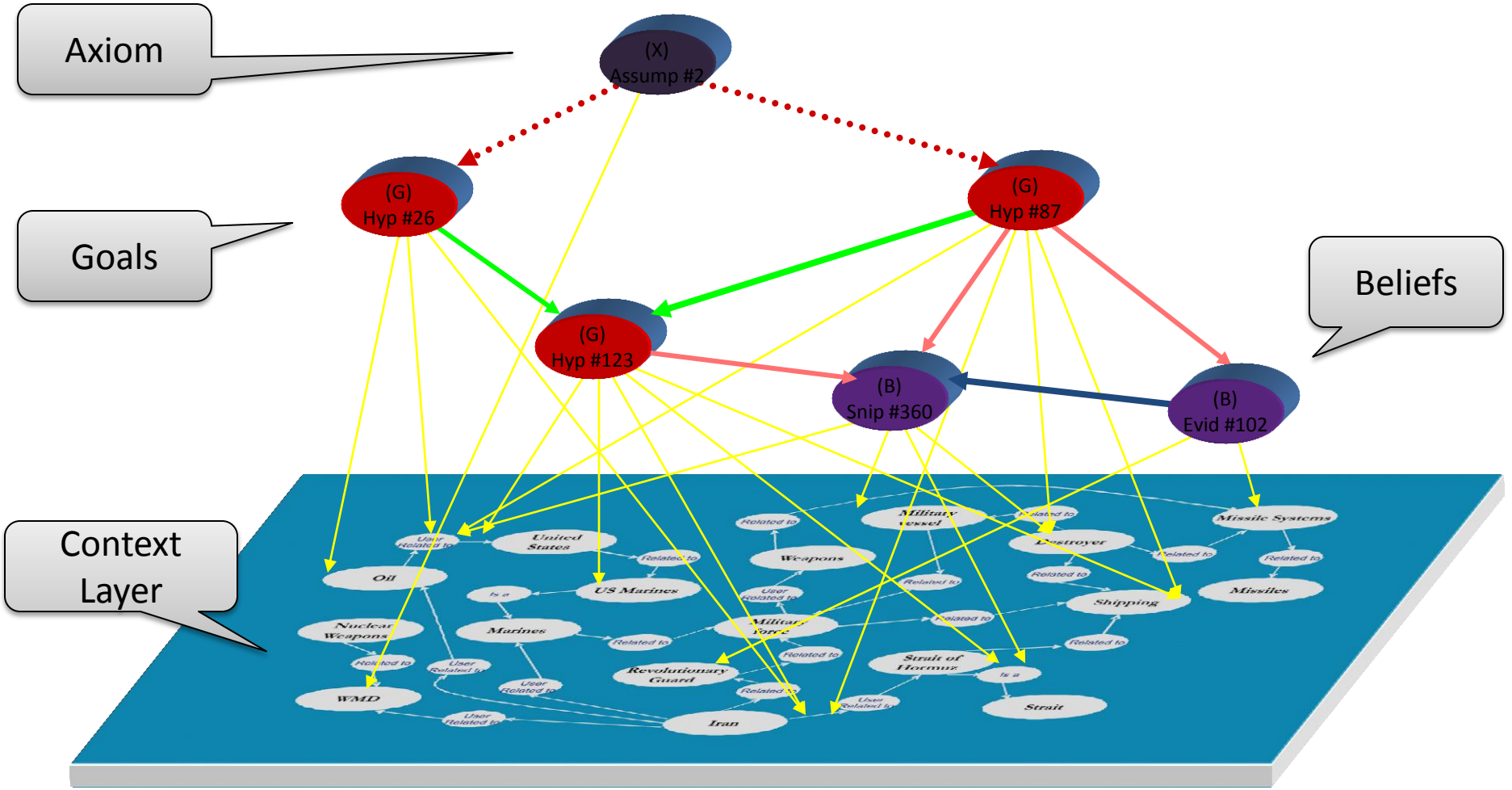
Rationale Network

- Encodes the high level goals of the insider, beliefs that support these goals, and the context within which these beliefs are held.
- Typology of Nodes –
  - **Context** – Concepts and relationships among them.
  - **Beliefs** – what the insider believes about something or in something based only on collected/gathered info
  - **Goals** – what the insider is aiming for or trying to reach/prove
    - Hypothesis – what the analyst is trying to "support" or "prove"
  - **Axioms** – what the insider believes in not based on collected info
    - Intelligence doctrine/training of the insider
    - Personal Beliefs

Rationale Network Construction

- Initialize the rationale network with a goal taxonomy
- Define a set of actions that are pertinent to each goal
- When a new action occurs:
  - If a goal directly or indirectly related to that action is found in the goal taxonomy:
    - Generate a context network for the textual content associated with the given action.
    - Connect all the concept nodes generated to the chosen goal.
    - A belief node is added if a user makes explicit what he believes in (e.g., statements in an annotation).
  - If a goal corresponding to that action is not found in the goal taxonomy:
    - A new goal node is created and connected to the concept nodes of the context network corresponding to the action.
    - A new goal is also automatically created if the set of belief and context nodes of the existing goal is only covered by that goal at most *t*% of the time with *t* being the cutoff threshold

Rationale Network Example

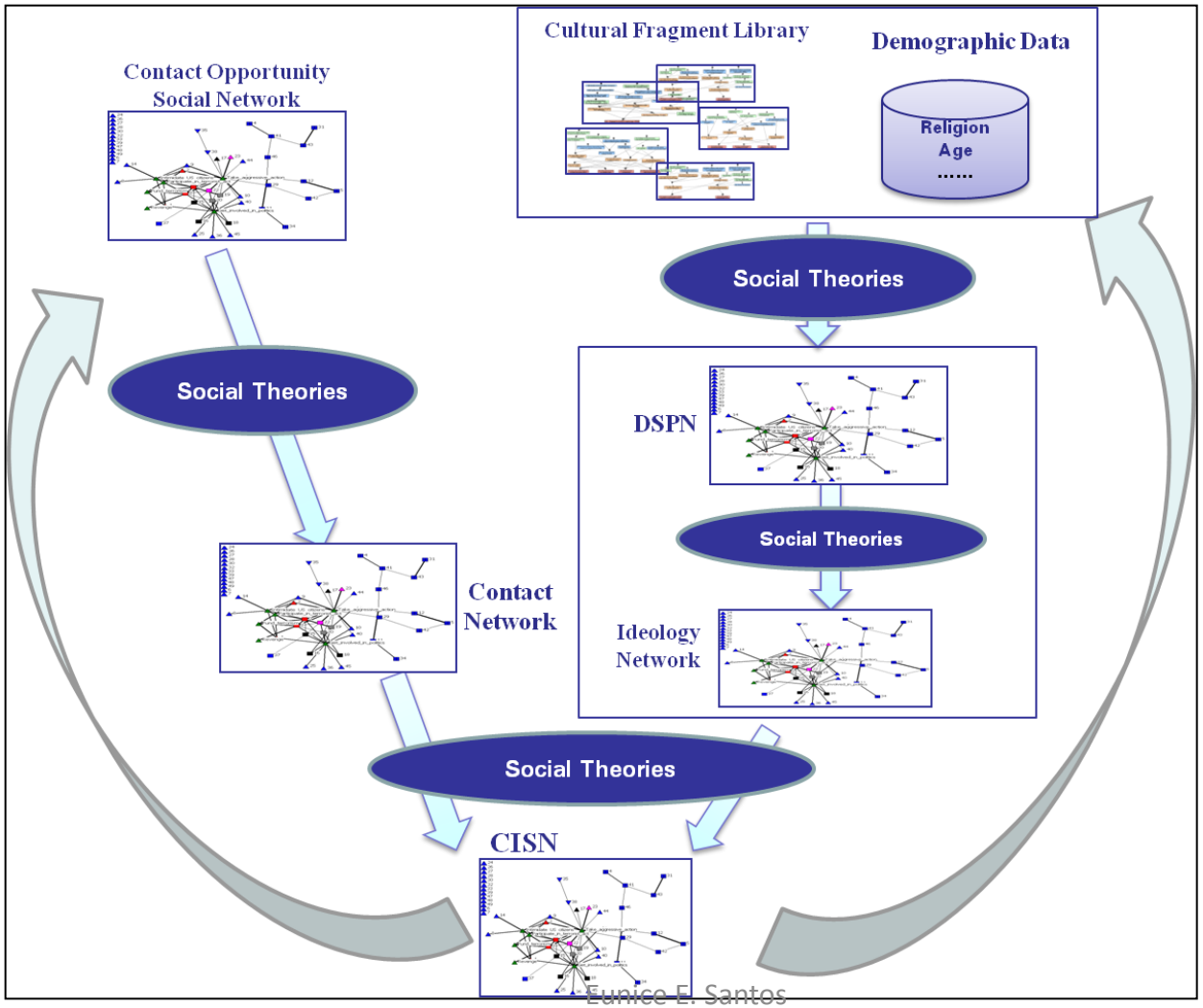


Foci Network

- Network used to track the foci of the analyst as they work.
- Each node is described by:
  - **Commitment Level** – how active the focus is.
    - $C(a) = \beta_f f(a) + \beta_f r(a)$ (where f(a) is frequency and r(a) is recency)
  - **Goal** – the particular goal relevant/discovered with respect to the focus
  - **Interests** – topical interests and level of emphasis that are relevant to the focus
- Each edge is described by the source and destination goal nodes and the type of link it represents.
  - *regular* links: represents the link between two goals as it is shown in the Rationale network.
  - *leakage* links: represents a relationship in which two goals have been fired together frequently.

- Divided into long term and short term foci

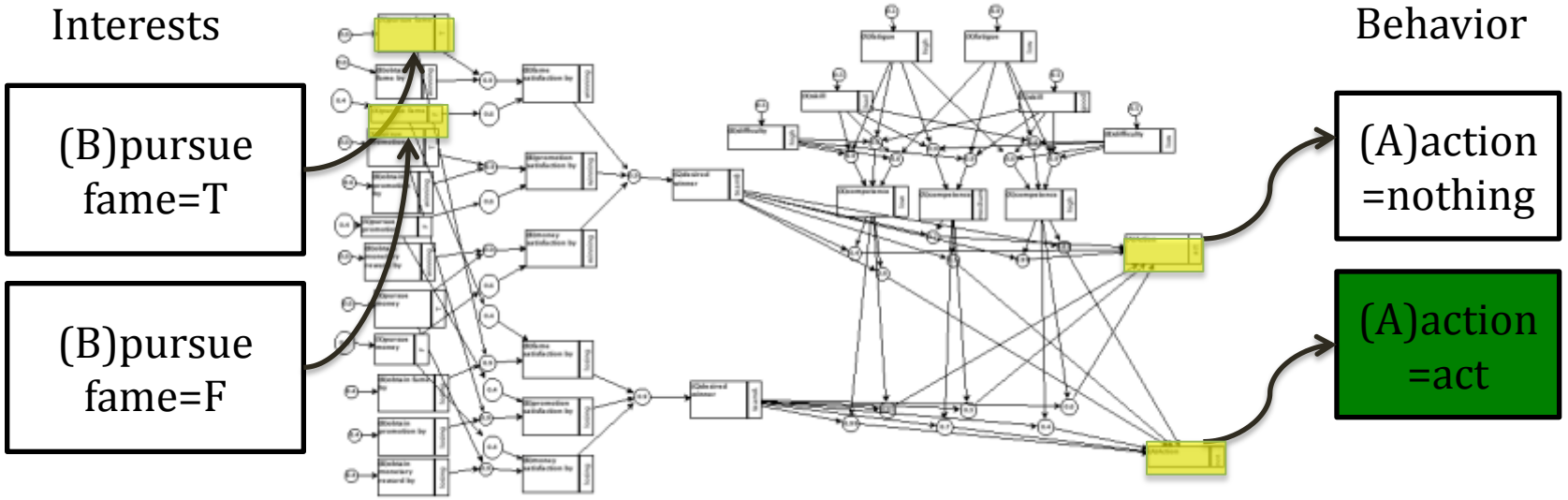
Culturally-Infused Social Network (Santos et al., 2008)



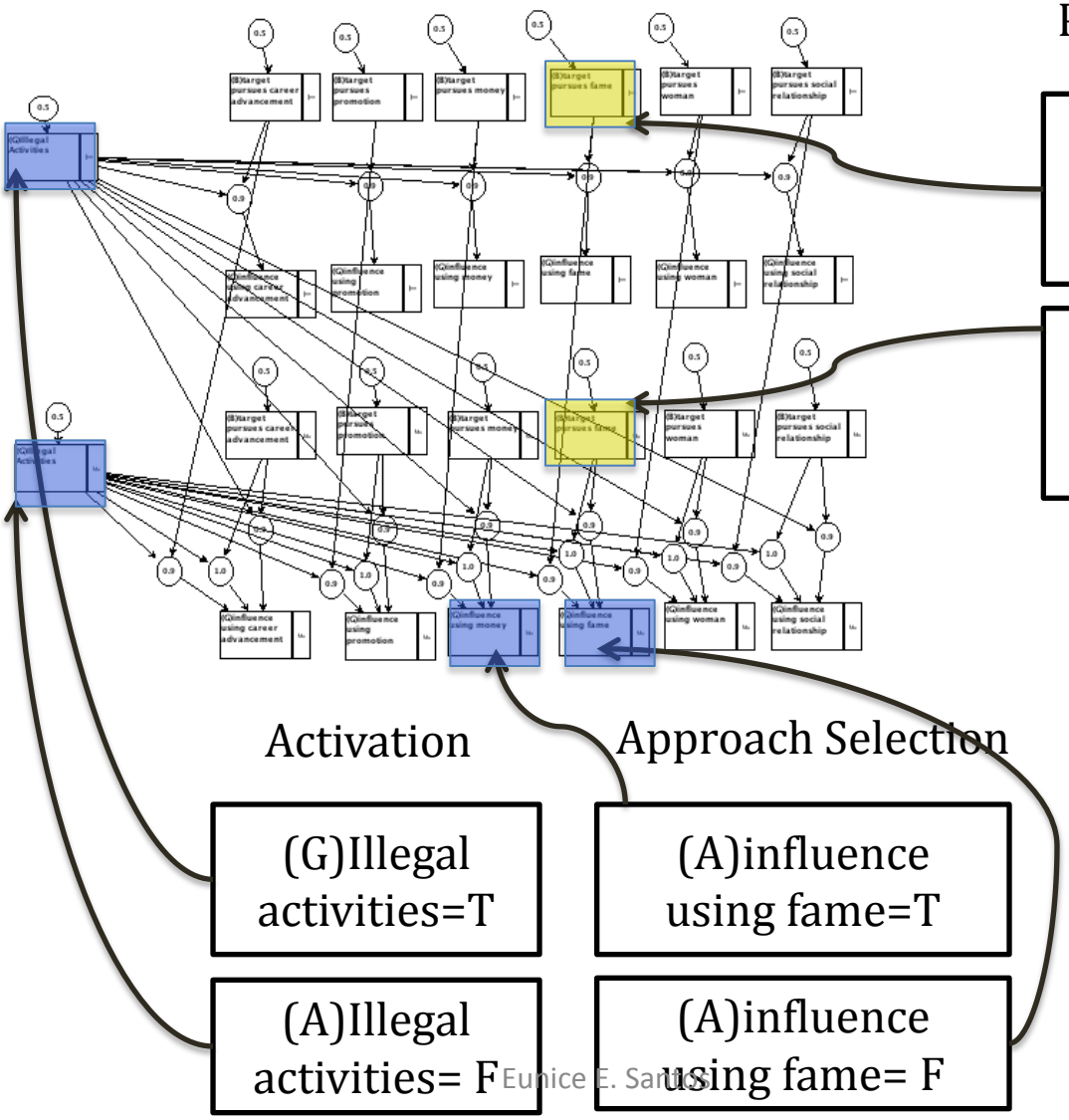
Methodology

- Modeling *behavioral change-targeted* interaction
  - LEVEL 1: Perception of the Influencee
    - interests
    - behavior/opinions
  - LEVEL 2: Influence Activation
    - Criteria 1: Motivation to communicate (culturally-constraint)
    - Criteria 2: Selection of offer to satisfy the interest of the influencee
  - LEVEL 3: Reaction to Influencer
    - evaluation of interests satisfaction
    - decide whether to accept the offer (leads to behavioral change or opinion change)

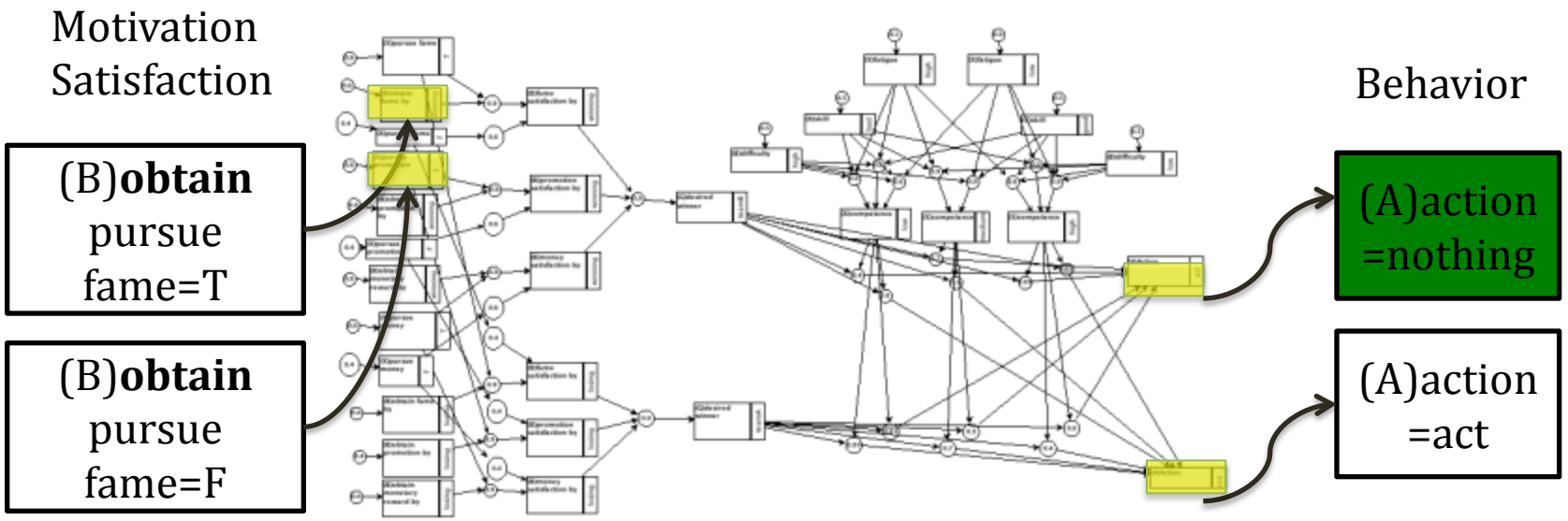
Level 1-Perception of Influencee



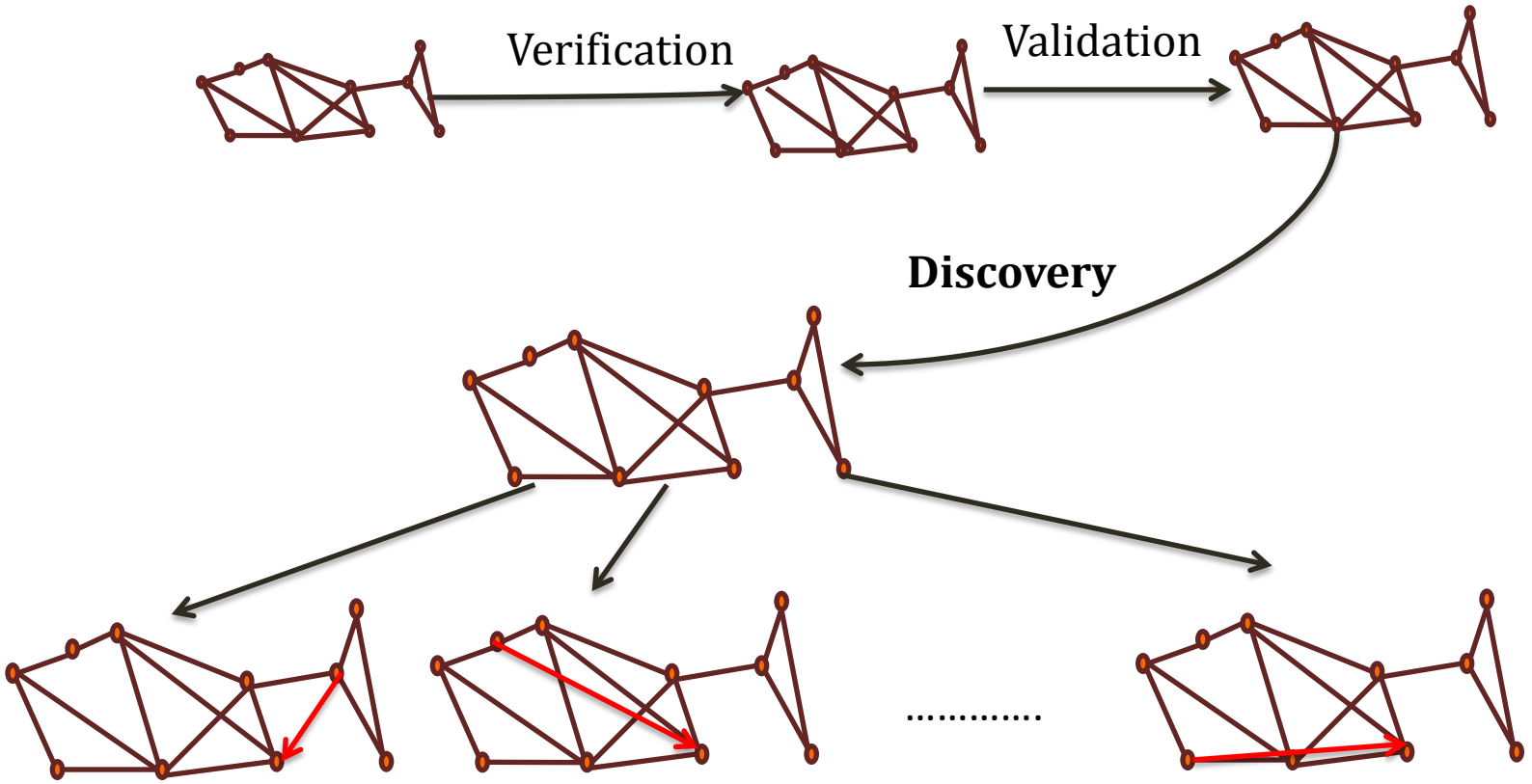
Level 2-Influence Activation



Level 3 – Reaction to Influence



Detection Methodology



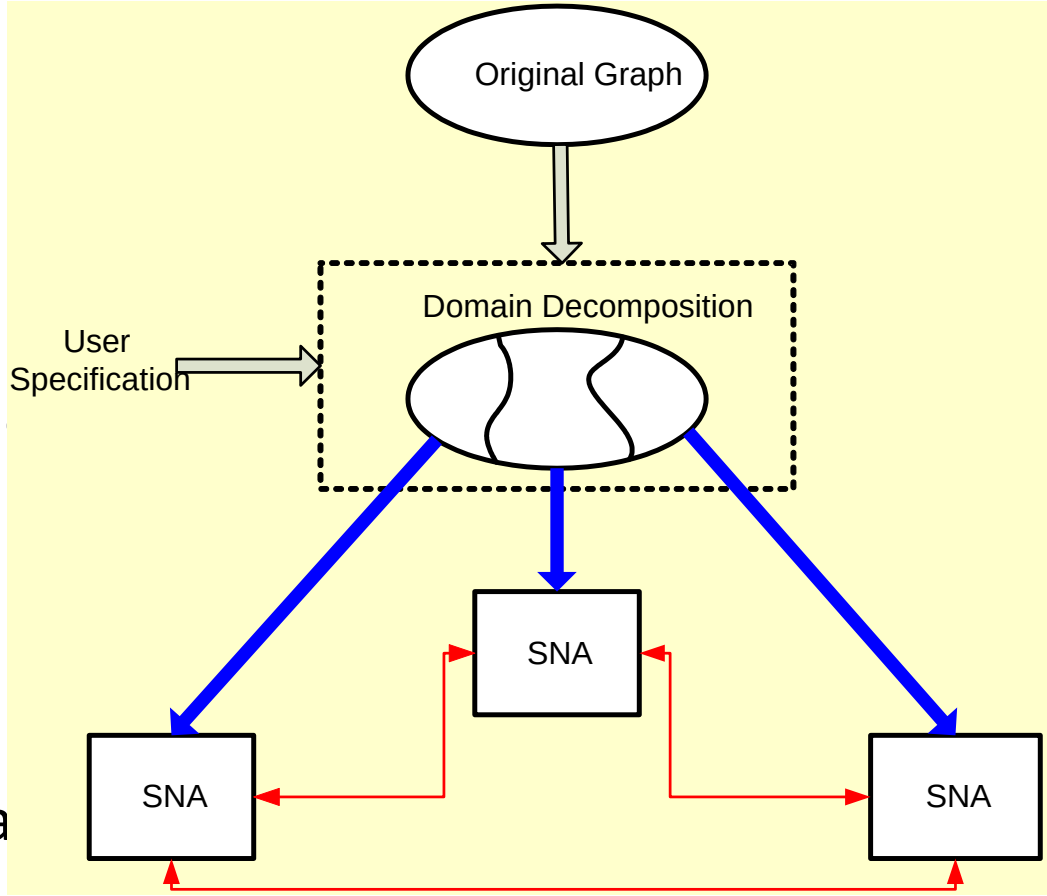
behavioral change
(a_2,a_1)
with confidence lv

behavioral change
(a_5,a_1)
with confidence lv

behavioral change
(a_n,a_1)
with confidence lv

Anytime Anywhere Methodologies

- Three Phases:
  - Domain Decomposition
  - Initial Approximati
  - Recombination
- Modular Design
  - One module for ea phase



Methodology Architecture [Santos PAP'06]

Take Away

- Cyber intelligence and operations will have a new capability to not only better catch malicious insiders, but be able to also understand why they occur, how they occur, and how they can be mitigated, managed, or manipulated.

Targeted Interventions Derived from Biomarkers of Cyber Trust

THRUST 2

Thrust 2 Team

- (Thrust PI:) Dr. John Hale, The University of Tulsa
- Dr. Rose Gamble, The University of Tulsa
- Dr. Bradley Brummel, The University of Tulsa
- Mr. David Greer, The University of Tulsa
- Dr. Patrick Bellgowan, Laureate Institute for Brain Research
- Dr. Jerzy Bodurka, Laureate Institute for Brain Research

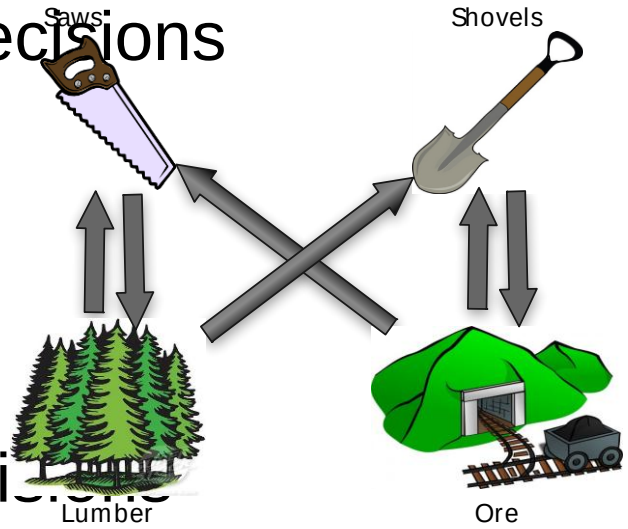
Research Problem and Approach

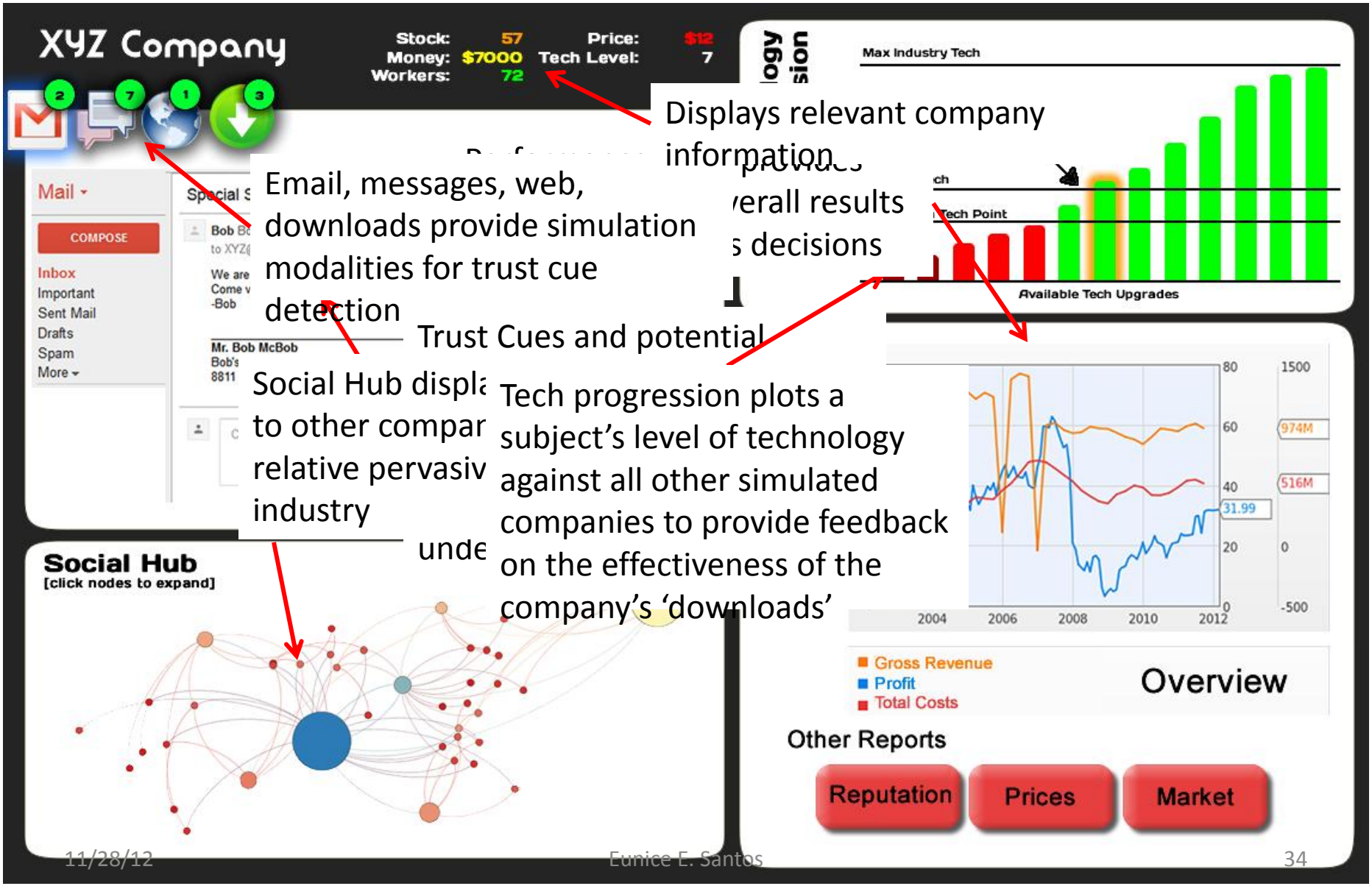
- Research Questions
  - Why and how do people make trust decisions online?
  - Can people be classified by the ways in which they trust?
  - Can classifications target more effective cyber trust training?
- 2 Phase Plan
  - Phase 1: Use fMRI + simulation to develop a classification map of cyber trust
  - Phase 2: Validate map and explore targeted training methods
- Contributions

1. Definition of the neural correlates of trust decisions in a cyber context.
2. Biomarker for cyber trust decision propensity.
3. Simulation platform for cyber trust research ("The Cyber Trust Game").
4. Evaluation of targeted interventions to mitigate trust errors.

Cyber Trust Game

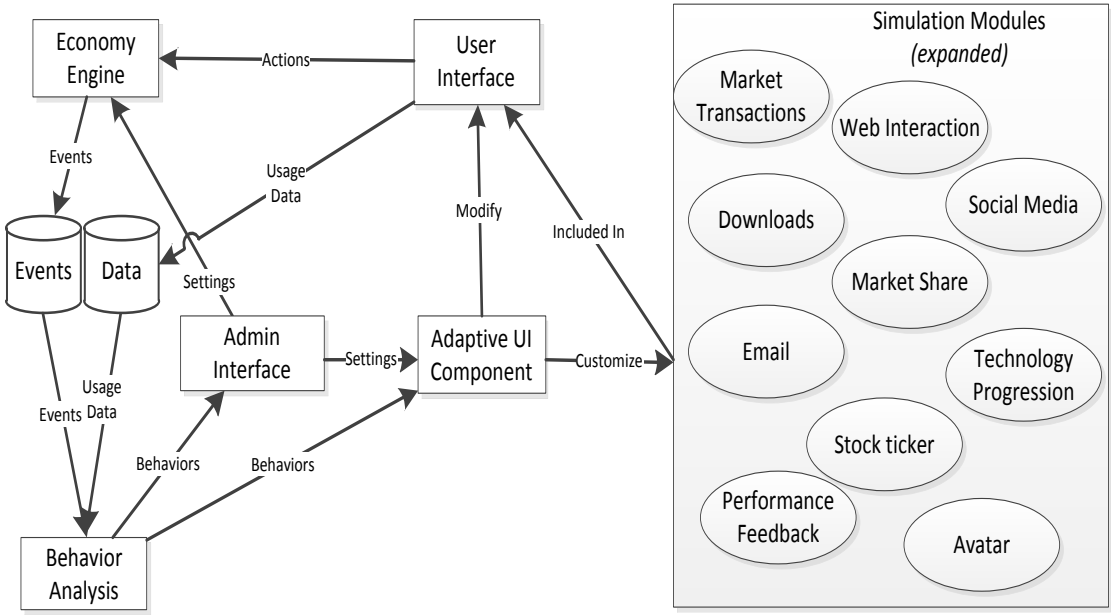
- Digital economy simulation
  - Closed economy B2B commerce simulation
  - Confront subjects with trust decisions
- Trust decisions
  - Email, Web, Social Networks
  - Trust cues
- Simple form play
  - Context free trust/no trust decisions
- Free play
  - Context dependent trust/no trust decisions




XYZ Company
Stock: 57
Money: $7000
Workers: 72
Price: $12
Tech Level: 7
Mail
COMPOSE
Inbox
Important
Sent Mail
Drafts
Spam
More
Max Industry Tech
Tech Point
Available Tech Upgrades
Displays relevant company information
Email, messages, web, downloads provide simulation modalities for trust cue detection
Trust Cues and potential
Social Hub displa
to other compar
relative pervasiv
industry
Tech progression plots a subject's level of technology against all other simulated companies to provide feedback on the effectiveness of the company's 'downloads'
Social Hub
[click nodes to expand]
2004
2006
2008
2010
2012
Gross Revenue
Profit
Total Costs
Overview
Other Reports
Reputation
Prices
Market

Application Architecture and Trust Cues

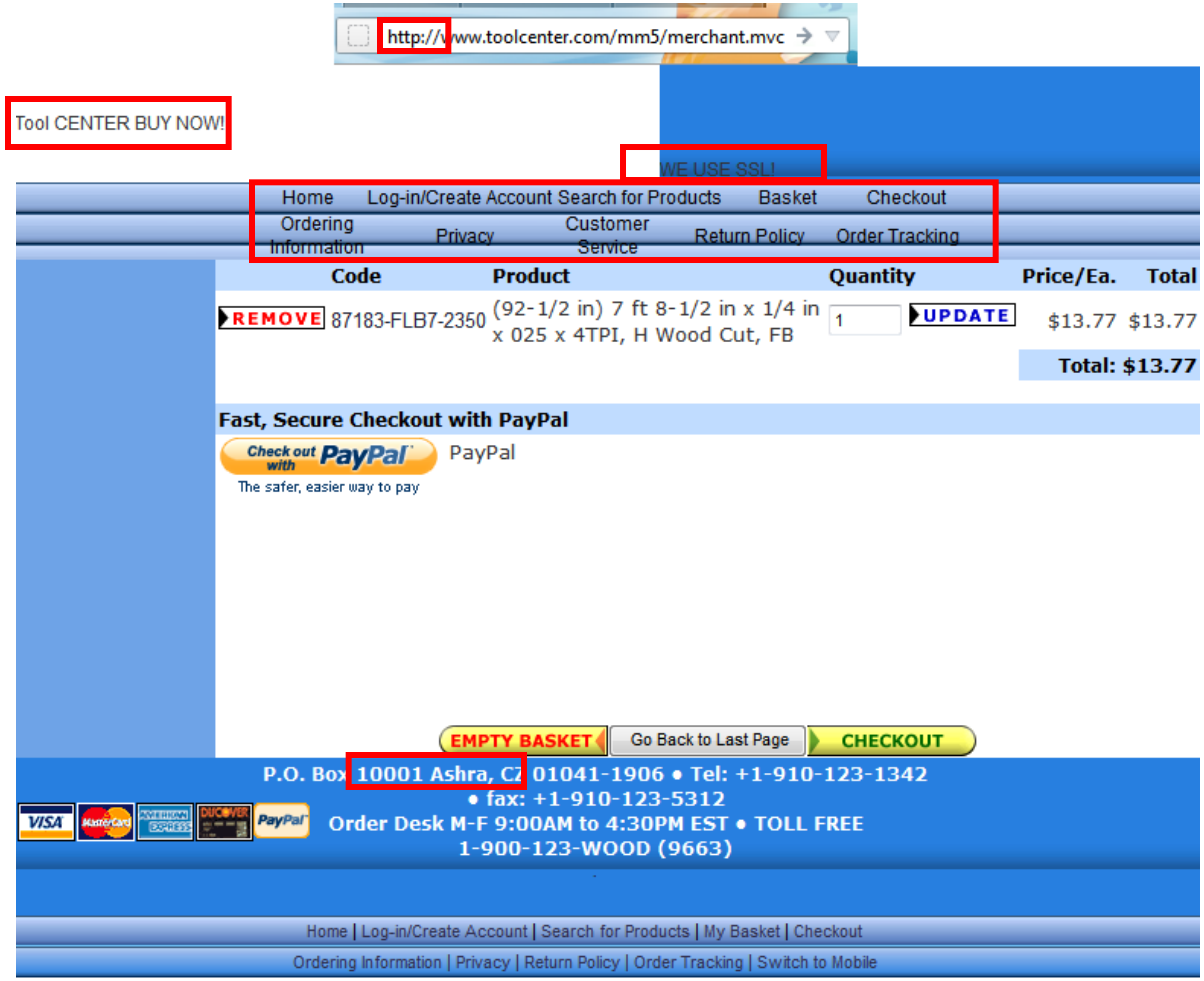
- Economy Engine – handles user actions and processes the corresponding events
- Behavior Analysis – identifies trust patterns based on the events and user interface usage data
- Admin interface – provides interface for viewing captured study data and making adjustments to the economic engine
- Adaptive UI – modifies trust cues based on study presets and user behaviors to understand how the modality and presentation of information affects the recognition of trust cues



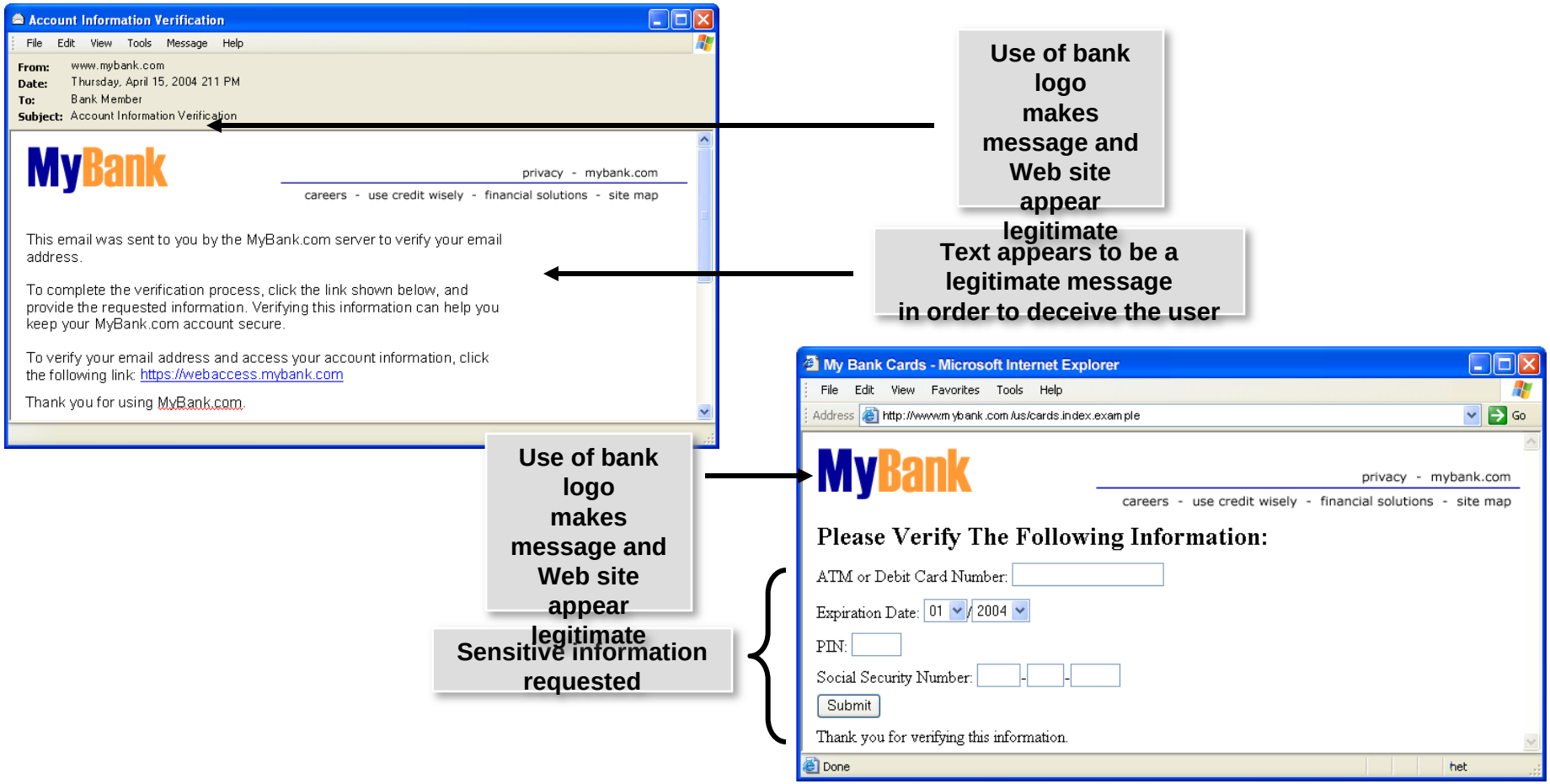
Modality	Trust Cues
Email	1 – Improper Email address, 2 – typographic errors, 3 link points to different URL, 4 – Generic greeting, 5 – Requests personal information, 6 – Contains unrealistic "get rich quick" claims, 7 – Source is outside of Home Country, 8 – Urgency of Email
Web	1 – Improper web address, 2 – typographic errors, 3 – Requests personal information / Confirmation of account details, 4 – Free offers, 5 – Source country code is outside of Home Country (e.g. .ru / .cz), 6 – Shopping / Web form is not protected by SSL (i.e. no "https" or "lock icon"), 7 – Excessive Advertisement, 8 – Improper session identifier, 9 – List of "search style" links to other sites, 10 Poor site design
Social	1 – typographic errors, 2 – Requests excessive access to personal information [app specific], 3 – Free offers, 4 – Requires sign-up for access [app specific], 5 – Has no Photo, 6 – "About section" is sparse, 7 – profile includes link to suspicious site, 8 – Excessive posting, 9 – No / Very little posting, 10 – Post contents are very commercialized, 11 – Low number of other friends/followers
Download	1 – suspicious source, 2 – Virus alerts, 3 – Free offers, 4 – Requires sign-up or subscription for access, 5 – Contains adds in the application, 6 – Prompts to install unwanted freeware (e.g. browser search bars), 7 – Downloading requires clicking through multiple links, 8 – Poor download design

Malicious Websites with Trust Cues

- Example Question:
  - Is this website trustworthy?
  - Answer: Yes / No
  - If No: What makes you believe it is untrustworthy? [Select all that apply]
  - Cues:
  - 1.Improper web address
  - 2.Typographic errors
  - 3.Requests personal information / confirmation of account details
  - 4.Free offers
  - 5.Source country code is outside of home country
  - 6.Shopping / web form is not protected by SSL
  - 7.Excessive advertisements
  - 8.Improper session identifier
  - 9.List of “search style” links to other sites
  - 10.Poor site design

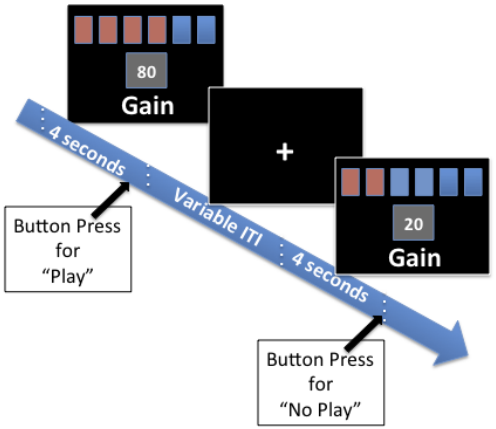


Phishing Email Example: Information Request



Functional MRI Network Localizer Tasks

Risk vs. Reward Network

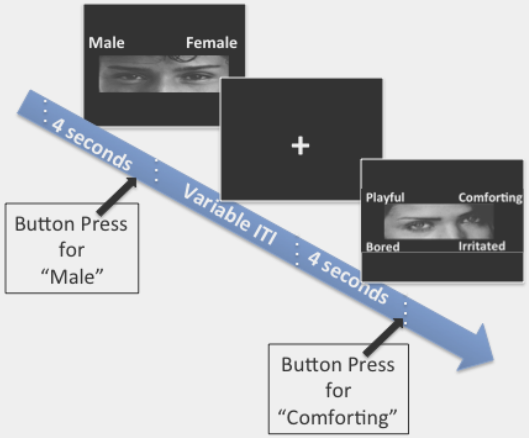


Task: Choose to Gamble or Not

Stimulus components:
1) Top bar depicts probability of winning

2) Number points at risk

Social Cognitive Network

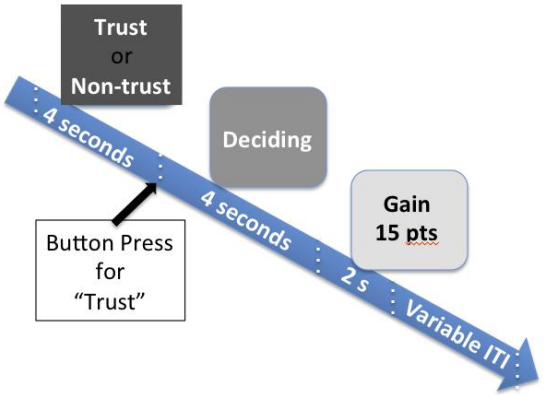


Task: Choose to which emotion Is being expressed by the eyes

Stimulus components:
1) Human face: eyes only

2) Possible emotions

Interpersonal Trust Network



Task: Cooperative exchange game.

Stimulus components:
1) Decision Screen

2) Exchange feedback.

Functional MRI Cyber-Trust Network Decomposition

Hypothesis:

Cyber Trust Network = Interpersonal Trust Network ⊗ Social Cognitive Network ⊗ Risk vs. Reward Network

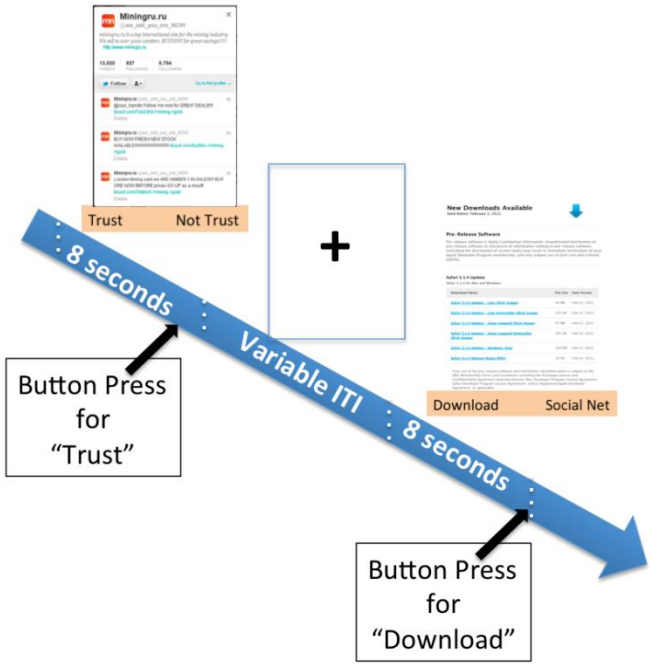
Cyber Trust Game Adapted for fMRI

Task: Choose to Accept or Not Accept the Cyber offer presented

Stimulus components:
1) Images of the various Cyber threats

2) Control conditions

3) Response Screen



Analyses:

1) fMRI network decomposition

2) Biomarker classifier

Eye Tracking for Evaluating Trust Cues

- Examining How Experts and Novices Read for Cyber Trust Cues
  - Are People Aware of What Cues Exist?
  - Do They Process the Cues When Making Trust Decisions?
  - Do the Cues Influence Decisions? For Whom?
  - What Makes a Communication Difficult to Discern its Trustworthiness?
- Formative Evaluation of Trust Cues
  - Allow for Reliable Measurement of Individual Trust Performance
  - Allow for Building Multiple Versions of Simulation Game

Training Intervention Design and Evaluation

- Adaptive Simulation Game
  - Practice Cyber Trust Decisions
    - With Real Time Feedback and Consequences
    - In a Realistic Environment that Encourages Engagement
  - Adapt Training to Individual Profiles
    - Awareness of Cues in Specific Modalities
    - Overly Heightened Trust or Suspicion Overall
- Training Evaluation
  - Does the Simulation Training Enhance Learning Beyond Classroom Awareness Training?
  - Can Elements of the Simulation Game be Distributed Widely and Efficiently for Cyber Security Training?

Goals

- Classification map for cyber trust
- Game simulation platform for assessment and training
  - Simple form and Free play versions
- Practical biomarker(s) for cyber trust
- Targeted intervention strategies and tools

A Human-Centric Approach to Cyber Trust and Suspicion

THRUST 3

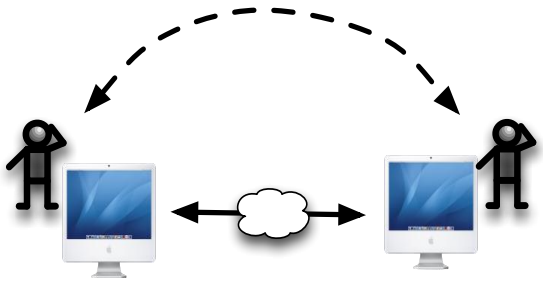
Thrust 3 Team

- (Thrust PI:) Dr. Hongbin Wang, University of Texas School of Biomedical Informatics at Houston

Why an empirical approach?

- **Cyberspace security is a technology issue as well as a human-social issue.** Unfortunately, the significant role of human factors in cyberspace security, including the processes and impacts of human operations, has not been fully recognized and understood.

- The goal of this research effort is to systematically demonstrate and examine how human performance in general and **human trust and suspicion** in particular may fundamentally affect cyber security operations with humans in the loop, and to explore how such effects can be mitigated or exploited in order to achieve a higher-level of security.



Research Objectives

- To empirically understand how human trust and suspicion in cyberspace are represented, measured, monitored and managed.
- To develop comprehensive computational theory and model of human trust and suspicion in cyberspace that can be compared and integrated with existing cyberspace technology so that new capabilities can be explored and implemented.

Research Questions

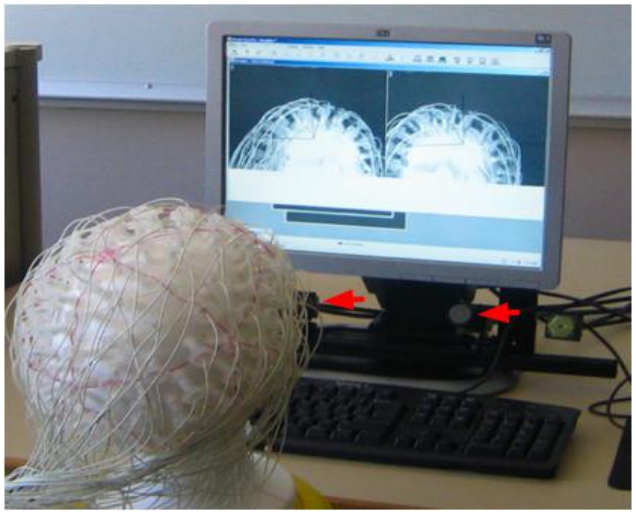
- What are cyber trust and suspicion (e.g., definition and taxonomy)?
- How are cyber trust and suspicion measured and indexed (e.g., theoretically, psychometrically, neurologically, and algorithmically)?
- How should people, and how do people, manage cyber trust and suspicion?
- How can human “biases” in cyber trust and suspicion be mitigated or exploited?
- Can we simulate the sensible (and insensible) human trust and suspicion behavior in cyberspace using an executable computational model? If so, can such a model be compared with, and integrated into, the traditional cyber-security models?

Technical Approach

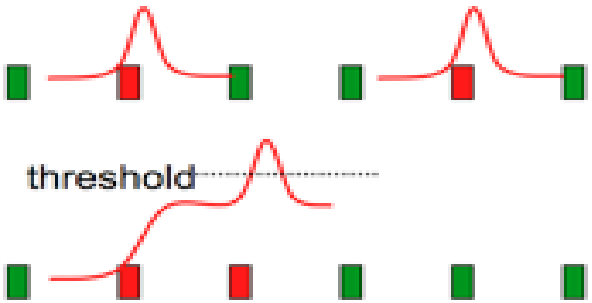
Theoretical Framework

- Trust and suspicion are loaded concepts with rich semantics.
  - Uncertainty
  - Confidence
  - Reliability
  - Credibility
  - Predictability
  - Benevolence
  - Emotion
  - Feeling
  - Vigilance
- In this project, we adopt an abduction-based framework and argue that trust and suspicion, with both symbolic and subsymbolic components, arise from a parallel constraint satisfaction process in a network that include all relevant observations, hypotheses and their relationships.

Neuropsychological Index



Integrated experimental system that combines behavioral, eye-tracking, and neuroimaging



Trust/Suspicion level indexed/monitored as events appearing

Using Non-invasive Sensors to Predict Trust and Suspicion in Human Operators

THRUST 4

Thrust 4 Team

- (Thrust PI:) Dr. Leanne Hirshfield, Syracuse University

Research Goal

- Run human subject studies (where subjects are equipped with non-invasive sensors) to provide real-time predictions about the changing level of trust and suspicion experienced by subjects while they conduct tasks that are designed specifically to test hypotheses stemming from the other team members' research.

AFOSR DURIP Funded Suite of Non-Invasive Sensors

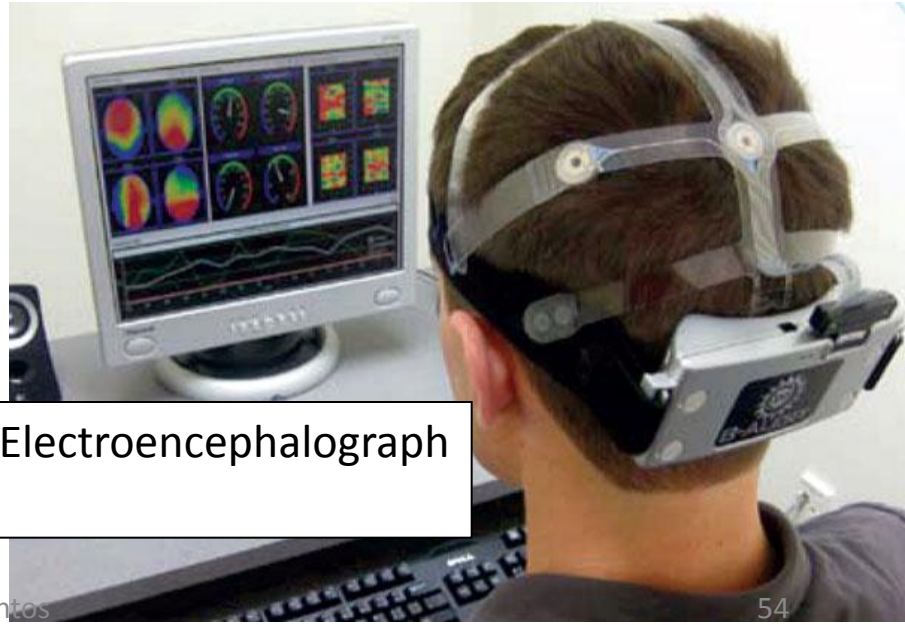
functional near-infrared spectroscopy

galvanic skin response

eyetracking

Morae
TechSmith

usability software



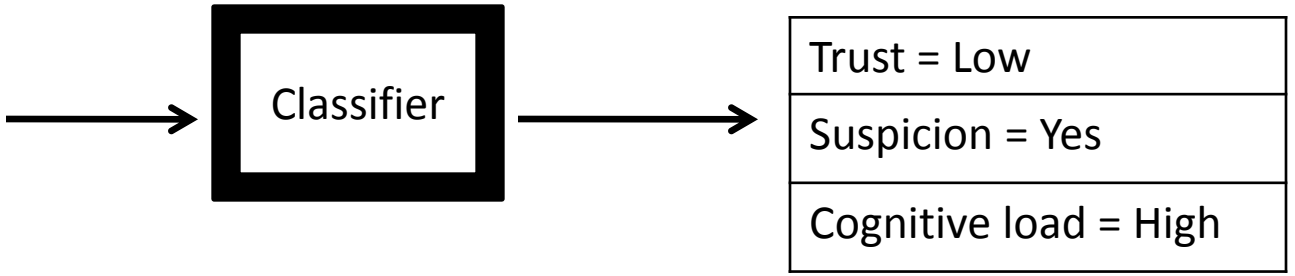
Electroencephalograph

Related esearch

- Define the constructs of trust, distrust, and suspicion in the IT domain.
- Use of non-invasive sensors to measure trust, distrust, and suspicion during realistic human-computer interactions.
- Refine definitions, experiment protocols, and machine learning techniques to ensure accurate, repeatable, predictions of trust, distrust, and suspicion under normal working conditions.
- Result of Hirshfield's Related Research: ***A 'trust classifier' that predicts trust and suspicion from non-invasive sensor data***

SU's Role within the Cyber Trust and Suspicion Team

Run human subject studies where the 'trust classifier' is used to test hypotheses from other team's research.



user works with IT system while wearing sensor(s)

time stamped data is sent to a machine learning classifier

a prediction is made about the user's level of trust at that time

SU's Role within the Cyber Trust and Suspicion Team

- The SU team will also work closely with all other groups to test hypotheses, design experiments, and to help create a cohesive tie between the research conducted by each group.

Planned Experiments With Each Group

- Thrust 1: SU will design and run experiments to test UTEP's models and hypotheses regarding insider threats.
- Thrust 2: SU will work with these researchers to compare the results between the MRI studies conducted in Tulsa, and results found when running the same study using the suite of non-invasive sensors at SU.
- Thrust 3: SU will design and run experiments to test the outputs generated by Dr. Wang's computer models.
- Thrust 5: SU will provide team AIS with keystroke and mouse data from human subject experiments to use in their analyses.

Assessing, Attributing, and Manipulating Operator Suspicion

THRUST 5

Thrust 5 Team

- (Thrust PI:) Dr. John S. Bay, AIS
- Mr. Robert Dora, AIS

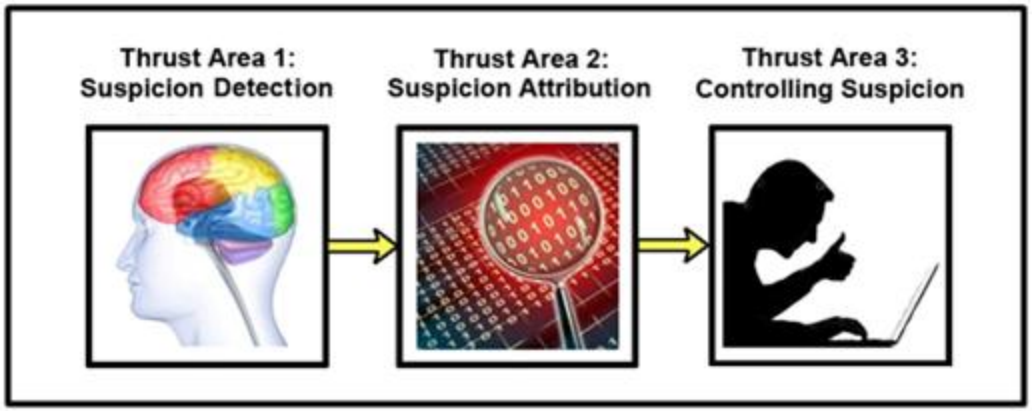
Current Research Tools

- **Interface Manipulation Platform (IMP)**
  - Command & Control (C2) application to launch interface manipulations
  - Primarily supports disrupt, deny, and deceive D5 effects
  - Currently integrated with many existing AIS, Inc. D5 effects, including:
    - Insert/drop keystrokes
    - Adjust screen flicker rate
    - Random mouse movements
  - Used at Hamilton College Next-Generation Usability Lab for the Deny and Disrupt effort
  - To be adapted and used during research thrust area 2
- **Remote Suspicion Identification (RSID) Keylogger**
  - State-of-the-art keylogging software developed under the RSID program
  - Captures keystroke timings and characters
  - Extracts and calculates key hold time, key interval time, key press latency, and key release latency
  - Consists of algorithms for calculating changes in keystroke patterns
  - Capable of capturing application focus and mouse movements

Technical Approach

- **AIS, Inc. will:**
  - investigate the detection, attribution, and manipulation of suspicion in users by means of non-invasive cyber sensors.
  - build on prior research in suspicion and extend the current understanding of suspicion among operators engaged in cyber activities.
  - investigate the impact of Cyber D5 (deceive, deny, disrupt, degrade, and destroy) effects on mental state.

- The effort will be spread across three years and three distinct research thrust areas: suspicion detection, suspicion attribution, & controlling suspicion.



Suspicion Detection (Year 1)

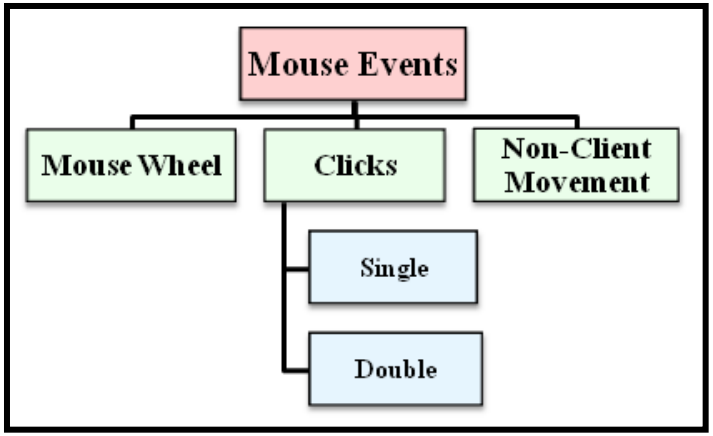
Keystroke Dynamics

- Correlation has been found between keystroke timings and changes to mental state, such as cognitive workload and deception under the Deny and Disrupt (DnD) effort
  - Traditional Timing Features
    - Key Hold Time (KHT) – Keystroke duration (aka dwell time)
    - Key Interval Time (KIT) – Time between the release of one key and the press of another (aka flight time)
    - Key Press Latency (KPL)
    - Key Release Latency (KRL)
  - User Features
    - Frequency of errors
    - Use of numpad
    - Use of shift keys (order and which shift key)
    - Use of shortcut keys

Suspicion Detection

Mouse Dynamics

- Investigate features from past mouse dynamics research for applicability to mental state
  - Pusara & Brodley (2004) calculated: distance, angle, and speed for selected pairs of points within temporal windows of data.
  - Schulz (2006) examined features of curves within mouse movement (e.g. curve length, number of points within curvature area, and inflection points) and computed a histogram of typical mouse movement curves for each user.
  - Ahmed & Traore (2007) used mouse movement, drag & drop, point & click, and silence (non-movement) for a histogram.
    - Calculated traveled distance, action type, movement direction, average movement speed, movement speed versus travelled distance, and time elapsed during movement
  - Feher, et al. (2012) created hierarchy from individual mouse movements to elaborate sequences and calculated "trajectory center of mass" and "third and fourth" moment.



Pusara & Brodley (2004) classify mouse data into a hierarchy of mouse events. Non-client movement refers to movement within an applications title and menu bars.

Suspicion Detection

Other Cyber Sensors

- **Investigate other potential Cyber Sensors**
  - *User Preferences Sensor*
    - Application usage profile
    - Usage time
    - Login times
    - Perform anomaly detection
  - *System Call Monitors*
    - Monitor system calls/other low-level system APIs
    - Monitor registry access
    - Determine users behavior in response to a change in mental state or the occurrence of a D5 effect
    - Profile user: level of knowledge, technical sophistication, etc.
  - *Application-specific Sensors*
    - Determine which buttons are pressed in a GUI
    - Identify specific menu options utilized
    - Popular and technically informative applications
      - Windows Task Manager
      - Microsoft Word

Suspicion Detection

Human Experimentation

- **Search Engine Experiment**
  - AIS, Inc. will support in a human subject experiment based on a previous search engine task experiment designed and performed by Dr. Hirshfield.
  - Subjects will locate various items via a search engine, building up to a fake website on the fifth and final day of the experiment that produces pop-ups to induce suspicion.
  - Subjects will be equipped with fNIRS and GSR to provide ground-truth
  - Computer terminals will be equipped with Cyber Sensors to capture user behavior
  - Correlation analysis will be performed on the data to determine relationships between digital data and mental state
  - Cyber Sensors will be updated based on the results of the experiment

Suspicion Attribution (Year 2)

- **Thrust Area 2 (Year 2) focuses on Suspicion Attribution**

Tentative Thrust Area 2 Goals:

- Research and develop new Cyber Sensors to determine the cause of a suspicious state
- Optimize, refine, and potentially re-purposed existing Cyber Sensors (developed during Thrust 1) to serve as or support Attribution Sensors
- Apply Cyber D5 Effects to Cyber Sensors and mental states to serve as the primary sources of suspicion
- Conduct Human Subject Experiment (#2) that induces D5 effects on subjects and gathers physiological (ground-truth) and digital data on mental state
- Analyze experiment data to correlate sources of suspicion with user behavior; unique behavioral patterns that correspond to particular suspicion sources will be identified

Controlling Suspicion (Year 3)

- **Thrust Area 3 (Year 3) focuses on Manipulating Suspicion**

Tentative Thrust Area 3 Goals:

- Research and develop new Cyber Sensors to guide operators into desired mental states
- Modify existing Cyber D5 Effects to induce specific mental states
- Develop new Cyber D5 Effects to induce specific mental states
- Identify ideal configurations for D5 Effects
- Develop an "Operator Mapping" that identifies the relationship between mental states, D5 Effects, and anticipated resulting actions
- Support and analyze the results of a Human Subject Experiment (#3) to help define and refine the Operator Mapping

Short-term Goals

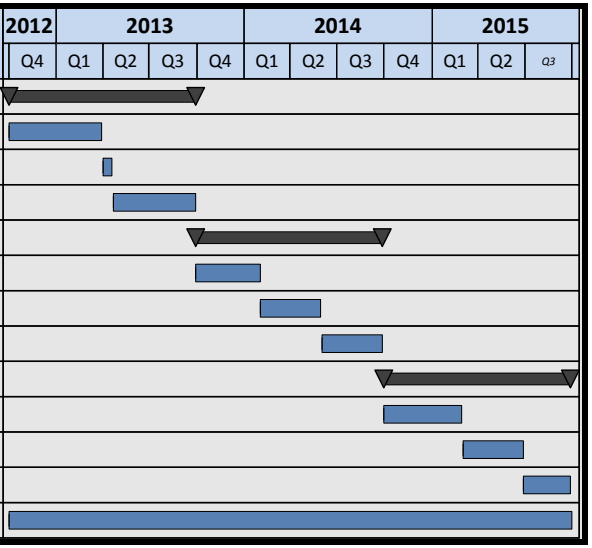
- **Thrust Goals:**
  - Develop Keylogging Cyber Sensor
  - Develop Mouse logging Cyber Sensor
  - Develop at least one other Cyber Sensor
  - Generate a digital behavior and physiological ground-truth dataset
  - Identify correlations in Sensor results with ground-truth experiment data
  - Modify Cyber Sensors to return specific mental states

Provide discreet and remote methods for detecting changes in operator mental state

Long-term Goals

- **Long-term Goals of the Cyber Trust and Suspicion Effort are:**
  - Generate discreet and remote methods for detecting changes in operator mental state
  - Generate methods that can assess the cause of any suspicion
  - Provide technology that allows for repeatable trials involving inducing D5 effects and measuring suspicion
  - Generate methods that can manipulate an operator's suspicion into a specific state (desirable or undesirable)

ID	Task Name	Start	Finish
1	**Thrust 1**	**10/1/2012**	**9/27/2013**
2	Cyber Sensor Development	10/1/2012	3/29/2013
3	Experiment Apparatus Integration	4/1/2013	4/19/2013
4	Suspicion Detection Correlation Analysis	4/22/2013	9/27/2013
5	**Thrust 2**	**9/30/2013**	**9/26/2014**
6	Adapt Interface Manipulation Platform	9/30/2013	1/31/2014
7	Suspicion Source Correlation Analysis	2/3/2014	5/30/2014
8	Attribution Sensor Development	6/2/2014	9/26/2014
9	**Thrust 3**	**9/29/2014**	**9/25/2015**
10	Cyber Sensor Integration & Logic	9/29/2014	2/27/2015
11	Operator Mapping	3/2/2015	6/26/2015
12	Operator Mapping Data Analysis	6/29/2015	9/25/2015
13	Documentation & Closeout	10/1/2012	9/30/2015



Overall Project

- **Start date: September 30, 2012**
- **Each thrust provides unique research, design, analysis and/or development capabilities critical to the area of cyber-trust and suspicion**
- **Integration of capabilities and results to be performed through the duration of the project**