Assured Information Security Capability Briefing

Founded in 2001 and headquartered in Rome, New York, Assured Information Security, Inc. (AIS) specializes in high-risk research and development for the Department of Defense and Intelligence Community. AIS expertise spans artificial intelligence, machine learning, biometrics, cyber security, cyber operations, signals intelligence, audio analysis and exploitation, and other areas. AIS plays a leading role supporting the Federal Government in technology development, from fundamental research through sustainment, as well as in rapid R&D in support of urgent operational requirements.

AIS capabilities in AI/ML include advances in deep recurrent neural network technologies, reinforcement learning, transfer learning, explainability, natural language processing, audio analysis and exploitation, and modeling and simulation technologies. AIS developed the current state of the art biometric technique for user verification through keystroke dynamics. The team also provides leading expertise in user behavior and movement analysis. This includes development of user identification techniques through gait detection, which employ advanced feature selection methods to isolate anomalies in movement behavior at the individual level. It also includes activity detection techniques, demonstrating state of the art algorithms for identifying user activities (e.g., walking, running, cycling) via commodity mobile device sensors. These algorithms formed the basis for further research in detection of user movement anomalies consistent with infectious disease or traumatic brain injury. The AIS team brings further experience in user movement modeling and accuracy enhancement through ensemble development, drawing upon multiple modalities and algorithms to dramatically improve anomaly detection and recognize baseline, normal behavior. This research supports applications to broader and more general modeling techniques for human movement and activity. AIS also has experience integrating AI/ML technologies into traditional modeling and simulation platforms. This work centers on multi-domain military operations planning, including the DARPA SAFE-SiM integrated M&S platform, AFSiM, and other environments. In these applications, AIS researchers develop AI/ML technologies to probe model properties through automated scenario construction and execution, perform anomaly detection, guide model modification, and support analyst planning.

Consisting of more than 200 research scientists, engineers, and support staff, AIS possess a deep subject matter expertise across AI/ML, biometrics, movement modeling, modeling and simulation, and numerous other areas. AIS also hosts facilities to conduct classified research and development, with support for storage and processing of data up to and including TS/SCI. As a performer on more than 17 DARPA efforts and 2 IARPA efforts in the past six years, AIS researchers understand high risk, high reward research and development and achieved significant advances to the current state of the art under these efforts.

Valuation of Information for Covert Collection Computation and Transmission from Offender Red Systems (VIC3TORS)

VIC3TORS, an AIS led project under a larger DARPA/I2O program, involved the research and development of techniques to monitor adversary cyber actors. This included the application of AI/ML techniques used to track actors across numerous devices and modalities including gait.

Devices included traditional desktop systems, laptops, tablets, and mobile platforms. Modalities included accelerometer and gyroscope, keystroke data, mouse data, touchscreen interactions, and other information. The technology tracks users through biometric features, including keystroke biometrics, mouse movement behavior, and gait detection. The research resulted in a full order of magnitude improvement over the state of the art for keystroke biometrics, drawing extensively upon deep recurrent neural networks and other techniques. Furthermore, a portion of VIC3TORS focused on research into the identification of individual cyber operators based on their behaviors on target systems. The team collected data from 104 subjects all executing the same mission plan against the same target network. The users all performed network discovery, vulnerability probing, exploit delivery, information discovery, and exfiltration. The plan was bounded by features analogous to MITRE ATT&CK attribution descriptors for actor groups (e.g., objective, availability of a small set of exploits). Analysis of the data showed that, with similar TTPs at the group level, individual variance at persona-level existed in attribution features, such as commands and switches used, command patterns, and inefficiencies.

As a result of this research, AIS developed technology can not only detect realistic biometric and movement patterns in a population, but isolate anomalies specific to an individual. HAYSTAC can benefit from these models and expertise to support user movement modeling.

Generative Adversarial Network Spoofer (GANSpoofer) IR&D

GANSpoofer proved the feasibility of adversarial learning techniques for defeating behavioral biometric techniques, demonstrated against keystroke-based user verification techniques used in existing insider threat detection products, and more broadly applicable to other modalities (e.g., mouse, gait). The approach uses a specific type of generative adversarial network (GAN), referred to as a U-Net, which consists of a unique, 11-layer model that increases fidelity and accuracy. It uses a discriminator based on the VIC3TORS method to distinguish between input consistent or inconsistent with a given user. A generator is then trained to produce realistic outputs, for users in general and a specific, target user. The trained generator can then be fed arbitrary character strings and generate both keystrokes and timing data consistent with a specific user.

GANSpoofer was trained on 12 models across three users, varying the number of samples in training from .25% to 50% of the total sample size for that user. It was then tested on 5000 excerpted samples from multiple users. Inputs were used to generate similarity metrics for both training and test data, with the average training signature score of 82.67 and average testing signature average score of 70.07, indicating high degree of confidence in signature match. The resulting GANSpoofer technology allows users to perform actions on systems consistent with an arbitrary, trusted user who is monitored using behavioral biometrics.

The GANSpoofer effort demonstrates the inverse of the VIC3TORs biometric identification problem. We’ve shown that we can both detect the unique anomalies associated with an individual’s biometric behaviors and use this information to transform data into, not only realistic patterns at a population level, but patterns specific to that individual.

Warfighter Analytics using Smartphones for Health (WASH)

Under this effort, AIS developed an advanced deep learning model to classify physical activities with respect to phone placement (e.g., pocket, table, bag, etc.) with a high degree of accuracy (greater than 90%). The objective of the WASH program is to develop technology to analyze mobile phone sensors (e.g., accelerometer, gyroscope, etc.) to detect symptoms of TBI and infectious disease. The research focused on detecting minor deviations that yield digital biomarkers in which subtle variations can predict injury and disease symptomology. Under this effort, AIS developed an advanced deep learning model to classify physical activities with respect to phone placement (e.g., pocket, table, bag, etc.) with a high degree of accuracy (greater than 90%). This model is used to classify collected data and develop additional models for detection of TBI and infectious disease.

The approach used a GAN architecture to generate robust, realistic data and as a mechanism to triage data without the need for labels. This method allowed for the automatic identification of data with sufficient signal and could be used to verify the individual's biometric signature. Deviations from this signature could be used to recognize the user, while detecting differences attributable to TBI and infectious disease.

Cooperative Robots: Autonomous and Deep Learning Enabled (CRADLE)

Under CRADLE, AIS researched imitating the behavior of a human using a low-SWaP device to navigate between two points. Using observations from a human performing the task, a policy was trained to mimic the behavior while reducing the computation needed to execute the policy. Additional research was conducted to transfer policies from a simulation environment to the real-world where the cost of data acquisition is lower. CRADLE used the TurtleBot3 Burger for its research on deploying machine learning to low-SWaP devices. The TurtleBot3 is a small mobile robot that can be programmed through the Robot Operating System (ROS). ROS is installed on a Raspberry Pi, a small low-cost computer included with the TurtleBot. ROS provides an interface for controlling the robot and consuming sensor data (i.e., the camera and the Light Detection and Ranging (LiDAR)). The AIS team built a policy using a convolutional neural network (CNN) in the TensforFlow framework. The model predicts the velocity and the required turning angle for the TurtleBot based on the image the camera captured. The successful policy to imitate the behavior of the operator allows the TurtleBot to navigate between the same two points without any human intervention. While the policy can be quantitatively evaluated with a validation dataset, the true test is for qualitative evaluation of the policy running on the TurtleBot. To reduce the real-world data requirements, AIS researched transferring policies from a simulation to the real world. This was done with NVIDIA's Imaginaire library, which provides unsupervised image-to-image translation between two domains. A MUNIT model from the Imaginaire library were trained on datasets from the Gazebo simulation and the TurtleBot performing in the real-world.

This effort gives AIS relevant experience in translating simulation data to a real-world application focused on replicating human control. The technology developed under CRADLE gives AIS an intimate understanding of how to improve experience gained in a simulation for transfer to the real world.

COunterfactual Demonstrations for EXplanation (CODEX)

CODEX is an explainability research effort to enable the practical application of Reinforcement Learning (RL)-based AI for widescale use in both military and commercial systems. The objective of the ongoing CODEX effort is to develop counterfactual examples for explaining Reinforcement Learning (RL) agents. The approach focuses on developing a world model representation of the target RL environment and then using the world-model to develop counterfactual trajectories intended to explain agent decisions. The world-model will be used to visually represent counterfactual trajectories and summarized with text to ease user information burden. The resulting technology will show a user both what the RL agent expects will happen after it makes a decision and develop counterfactual examples (i.e., "what ifs") to show a user what the RL agent expects would have happened had it made a different decision. CODEX trajectory presentation is not an inherently interpretable modeling mechanism of the policy. This makes CODEX a global, post-hoc explainability method intended to build user trust in policy decision making rather than building trust in individual predictions.

The methods for deriving the counter-factual examples from the RL policies within the simulation environment will allow AIS to generate realistic anomalous data without the need for real world collection. AIS has also researched and developed NLP capabilities to use in annotating simulation scenarios to automatically summarize and explain the actions taken by an autonomous agent within a simulation.

Common Unifying Representation Embeddings (CURE)

Under the ongoing CURE effort, AIS investigates and develops novel and effective transfer RL techniques based on world-models that learn to disentangle what is common and what is not between domains in the form of common embedding spaces. This approach was inspired by recent work in image and style transfer where learning these common embedding spaces has produced strikingly effective transfer and generalization of images from one style domain to another. Unlike traditional transfer RL approaches that attempt to learn how to generalize and adapt individual policies from one domain to another, CURE developed methods to learn disentangled embeddings, which are latent models of two domains that explicitly represent what is shared between the domains and what is not. By modeling what is shared between domains, instead of what is shared between policies, CURE enables broader and more effective transfer through improved domain adaptation and cross-domain planning. CURE uses these disentangled embeddings to develop transfer-aware policy learning algorithms that actively use knowledge of shared embeddings to effectively boost transfer and generalization across domains. Additionally, CURE uses the shared embeddings and examples of positive and negative transfer to train a metric model to predict transfer success. Such a metric provides a principled and effective method for determining the degree to which two domains can transfer policies from one to another.

The experience gained under the CURE effort gives AIS domain expertise in RL knowledge transfer techniques. These techniques can be applied to simulations environment for faster and intelligent transferring of policies (i.e., experience) across disparate domains.

Semantic Model, Annotation, and Reasoning Technologies (SMART)

The SMART effort is an ongoing applied research effort to develop a semi-automated framework to annotate simulation models for increased utility and efficiency of model composition. SMART will achieve this by combining assistive Natural Language Processing (NLP), automated model modification and behavior analysis, and a guided semi-automatic analyst annotation workflow. SMART combines assistive NLP, automated model modification and behavior analysis, and a guided semi-automatic analyst annotation workflow to quickly and semi-automatically annotate and compose simulation models from a variety of domains and at all levels of engagement (e.g., single strike or full campaign). This enables analysts to quickly identify models relevant to mission scenarios within a target security classification or under time and space constraints. SMART additionally addresses security classification tracking issues by providing semantic annotation of models and allowing characterization model security classification for a wide variety of model configurations.

Relevant to HAYSTAC, AIS has developed a comprehensive and detailed understanding of simulation models for complex domains using NLP and behavioral assessments. This program involved the development of an automated and intelligent parameter selection to enable model corrections.