**Research Activities and Plans**

1. **Intelligent Autonomous Systems based on Machine Learning and Data Analytics.**

**Problem Statement:**

Intelligent Autonomous Systems (IAS) are highly cognitive, reflective, multitask able, and effective in knowledge discovery. Examples of IAS include autonomous vehicles, network of sensors with reconfigurable sensory platforms, and an unmanned aerial vehicle respecting privacy by deciding to turn off its camera when pointing inside a private residence. Research is needed to build systems that can monitor, learn, understand its capabilities, and adapt to meet mission objectives with no human intervention. The system should be fail-safe and should allow for graceful degradations and continue to meet the mission objective. This leads to following research questions: (a) investigate cognitive processing approaches that enable the system to be aware of the underlying environment where the provenance data is being generated and (b) facilitate learning from historical data to improve performance. In addition, we need research in developing software to (c) interpret the high-level commands for real-time action and (d) contribute to representation and reasoning approaches based on semantic and statistics to discover new knowledge. Research is needed in (e) learning algorithms to enable autonomy to reason about making decisions under uncertainties.

**Proposed Approach****:**

We propose the following approaches to create autonomy in smart grids and contribute to “Smart Autonomy” NGC IRAD of Dr. Donald Steiner.

1. Employ machine learning analytics to learn and understand the underlying patterns in data provenance, conduct forensics to detect anomalies, and assist decision making by on-the-fly semantic and probabilistic reasoning.
2. Equipped with discovered knowledge, automate system’s self-awareness, self-healing, and self-optimization.
3. Utilize blockchain technology for storing provenance data for providing trust, immutability, and consensus and use NGC-WaxedPrune/Active Bundle [Bharat Bhargava, Donald Steiner, Jason Kobes et. al. IEEE Cloud 2017] to build a prototype.
4. **Cross Domain Information Sharing and Exchange:**

The CDIS Group addresses the growing war-fighter need to better share information appropriately across many organizations, including the coalition partners. The cross-domain Information sharing challenge is compounded by the need to transfer time-sensitive information across multiple security domains. The responsibility to share information must be balanced against the need to protect sensitive information. Joint work with Mike Mayhew and Mark Linderman at AFRL

**Problem Statement:**

The biggest threat to mission assurance is the lack of sharing of critical information in a timely and accurate manner in the cross-domain environment. This research will contribute to science of privacy, trust, and security in data dissemination among security domains.

We work towards a framework for de-centralized information sharing to overcome such problems, while considering multiple aspects of the framework including privacy, integrity, and trust. Our research requires bringing together expertise in databases, quality of service (QoS), privacy, trust, and context awareness/situation awareness.

The research involves the discovery, propagation, and aggregation of information shared by multiple participants across domains under varying situations and contexts. The system must adapt to the type, extent, duration, and timing of multiple attacks/failures. Our intellectual contributions include the development of algorithms for proactive dispersion of information, situation-aware paradigm, integrity checks and violator identification methods, information adaptability, and active bundles.

The thrust of this researchincludes dissemination of private data, privacy and trust, privacy metrics. Data dissemination should assure that different organizations can share their sensitive data without compromising privacy. Algorithms will be designed to evaluate the privacy loss due to disclosure of information and gain in trust. A series of experiments will provide guidelines in privacy measurement, trust assessment, and quantification of the tradeoff between privacy and trust. A privacy assessment metric will be developed that employs information theoretic approaches to measure privacy. Various privacy violator models and user behaviors will be used as benchmarks for testing and evaluating different privacy-preserving techniques.

**Proposed Approach**

We propose a novel solution to the problem of preserving privacy in data dissemination. It is based on three ideas: self-descriptivenessof private objects, object apoptosis(clean self-destruction), and proximity-based evaporationof objects. The new concepts are Active Bundle and Managed Information Object with meta data and policies. Whenever delivery of a completeobject fails, the receiving guardian can recover it easily by retransmission. With atomic self-descriptive objects, there is no way that a sending guardian can transmit to the receiving guardian an object that is incomplete. This is true for every link of a privacy dissemination chain, and solves the problem of preserving privacy in data dissemination.

**Self-descriptiveness:** Sensitive data is accompanied within private objects by their metadata. Comprehensive metadatashould include owner’s privacy preferences, guardian privacy policies, metadata access conditions, enforcement specifications, data provenance, mandatory access labels and other context-dependents.

**Apoptosis (Clean Self-destruction):** When an object is about to be compromised by an attacker or an accident, autonomous apoptosis mechanismwithin an object can be implemented as a set of detectors and triggers setting off associated apoptosis code. The code is activated whenever a credible threat of a successful attack on the object is determined by the detectors. Situations in which the self-destruction trigger is overly sensitive and causes premature “suicides” can be dealt with by privacy recovery. A guardian may be able to recover object from a guardian preceding it.

**Proximity-based Evaporation:** A subtle reduction of data over time or events to reduce its sensitivity or impacts of dissemination based on environmental factors. The concept is that appropriately injected noise will make data less sensitive and their owners less vulnerable. We want to investigate the idea of using evaporation in conjunction with the apoptosis mechanism. As an object drifts further away from its original guardian, its apoptosis trigger becomes more sensitive. Data distortion may include replacing exact data with approximate data, or up-to-date values with previous value. Unauthorized data disclosures become more probable as distance increases.

1. **Cloud Privacy and Security**

I just published a sensitive document on Cloud. How can I remove it? How can I prevent leakage of my data? How can I avoid sharing memory and other resources with an adversary (Multi-tenancy problems)? Other issues include loss of control, lack of trust (mechanisms) and identity management. See tutorial at <http://www.cs.purdue.edu/homes/bb/cloud.html>

The end to end security challenges in SOA and Cloud are as follows:

* Authentication and authorization may not take place across intended end points
* Intermediate steps of service execution might expose messages to hostile threats
* External services are not verified or validated dynamically (Uninformed selection of services by user)
* User has no control on external service invocation within an orchestration or through a service in another service domain
* Violations and malicious activities in a trusted service domain remain undetected

Research on Cloud computing for Blind: Context-awareness is a critical aspect of safe navigation, especially for the blind and visually-impaired in unfamiliar environments. Existing mobile devices for context-aware navigation fall short in many cases due to their dependence on specific infrastructure requirements as well as having limited access to resources that could provide a wealth of contextual clues. In this work, we propose a mobile-cloud collaborative approach for context-aware navigation, where we aim to exploit the computational power of resources made available by Cloud Computing providers as well as the wealth of location-specific resources available on the Internet to provide maximal context-awareness. The system architecture has the advantages of being extensible and having minimal infrastructural reliance, thus allowing for wide usability. A traffic lights detector was developed as an initial application component of the proposed system and experiments performed to test appropriateness for the real-time nature of the problem.

1. **Collaborative Attacks and Defense in Networks**

Information assurance in cyber systems is critical for non-stop operations that support the information and communication requirements that keep a soldier safe and at peak performance. Past efforts have dealt with dealing with individual attacks. Coordinated attacks can cause havoc for the databases and networks and are hard to anticipate, avoid, detect, and defeat. The problem is exuberated if the multiple attackers can coordinate and gain the knowledge from each other.

Such attacks may overlap and run concurrently, follow one after the other, attack during recovery, and corrupt a large part of database or network. The damage could be geographically distributed or be concentrated on a small part of critical cyber operations.

The next step in research is to identify the events in defense against individual and collaborative attacks. The defense strategies have to be fast, comprehensive, and immediate. They should act based on some minimal knowledge of attack events and any coordination activity among them. Even stopping one crucial event from happening may mitigate the whole attack.

The context in which an attack takes place could help us in mitigating the progress of attack with minimal effort. The study of context has been used in many applications and can be utilized in building adaptable systems that can respond to the timing, duration, extent, severity of attacker’s intent. Context may change over time and should be considered in any adaptable defensive strategies. The research questions that need investigation are: How can one derive context and what is the source of its origination, to what use can the context be put to use? Collaboration needs details of environment for the largest impact. Context of either the defense or attack can enhance the capabilities of attacker as well as defender.

**Problem Statement:**

Develop and experiment with algorithms for survivability and recovery that provide information assurance, integrity of data and communication, confidentiality, and reliability in the presence of coordinated attacks. Defense should adapt to the type, severity, extent, timing, and coordination of attacks. Investigate approaches that can deal with the context of the mission to be accomplished and reduce threat and vulnerabilities. The context of a dangerous situation can specify the tradeoff between urgency of availability of resources versus privacy policies, evaluating trust. See our papers on [www.cs.purdue.edu/homes/bb/#research](http://www.cs.purdue.edu/homes/bb/#research) under section on journals and conferences. Titles are a) Secure and Efficient Access to Outsourced Data, b) Developing Attack Defense Ideas for Ad Hoc Wireless Networks, and c) Defending against Collaborative Packet Drop Attacks on MANETs.

1. **Assured Execution ( Briefly)**

Attacks on the IT infrastructure keep growing in volume and sophistication. As reported by Symantec, the release of malicious code and other unwanted programs may be exceeding that of legitimate software applications. On the other hand, a trusted execution environment is essential to assuring dynamic mission objectives.

The objective of this research is to contribute to the development of a fully trusted execution environment through the enhancement of existing virtualization architectures. The proposed research involves the development of a virtualization architecture, which will provide a secure run-time environment by greatly reducing the size of the trusted computing base of the execution environment. A virtualization architecture prototype will be developed based on the ideas proposed for trusted execution and extensive experiments will be performed to test the effectiveness of the architecture against various attacks. The prototype will also be evaluated in terms of execution time performance and algorithms will be developed to ensure the optimal performance.

**Approach:**

The approach for a trusted execution environment relies on a virtualization-based architecture to provide enhanced security by abstraction of physical resources into many separate logical resources. The security shortcomings of existing virtualization architectures will be addressed with two main ideas: 1) Secure execution of guest virtual machines (VM) under an untrusted management VM and 2) Preserving the run-time integrity of the hypervisor through greater control.

Many of the existing hardware-based approaches for a trusted execution environment are centered on the availability of the Trusted Platform Module (TPM) in the computing environment. TPM is a computer chip (microcontroller) (available by both Intel and AMD) that can securely store artifacts such as passwords, certificated or encryption keys, used to authenticate the platform. A TPM provides mechanisms of integrity measurement, storage and reporting of a platform, from which strong protection capabilities and attestations can be achieved. Dynamic *root of trust* measurement (DRTM) (as well as static root of trust measurement, which is not acceptable due to lack of scalability) is a widely adopted technique to start a chain of trust. Initially, the CPU must be in a known state, running known code and the system is in a state in which the code can defend itself. From this condition, each of the state changes can be measured by TPM to make assertions about the state of the computer. DTRM allows for a secure launch, where the hypervisor comes up in a trusted state, with the control of the system, regardless of what code has run previously. The main problem with this approach is that the System Management Mode (SMM) code of the machine is loaded before the initialization of DTRM. One major requirement for an integrity-protected hypervisor is that it has control over SMM code. In today’s virtualization architectures, SMM has unrestricted access to all system resources such as critical CPU registers, memory, and device IO locations. A buggy or malicious SMM code can therefore access memory regions belonging to the hypervisor and compromise its integrity. It has been shown that Intel’s TXT, which has been widely adopted as a trusted execution environment solution, is vulnerable to such attacks. We propose to limit the access capabilities of the SMM by running it in a container controlled by the hypervisor. This innovative approach along with the control of the hypervisor over all virtual machines (including the management VM), as described in the next section, will allow the hypervisor to have full control over the run-time environment and preserve its integrity against attacks, leading to a fully trusted execution environment.

1. **Mobile Security ( Briefly)**

If a mobile device is lost or falls in wrong hands, the data on phone must be securely removed. If the user of phone is in allocation that is insecure, some of the sensitive data must be blocked from access to prevent inadvertent disclosure. Based on the location and other parameters obtained from built-in sensors in mobile phone, various levels of privacy can be ensured. We are developing ideas for data loss prevention with industry and our ideas on cross domain information sharing, context awareness will be applicable. Our earlier work is secure routing, identification of intruders, and authentication while moving.