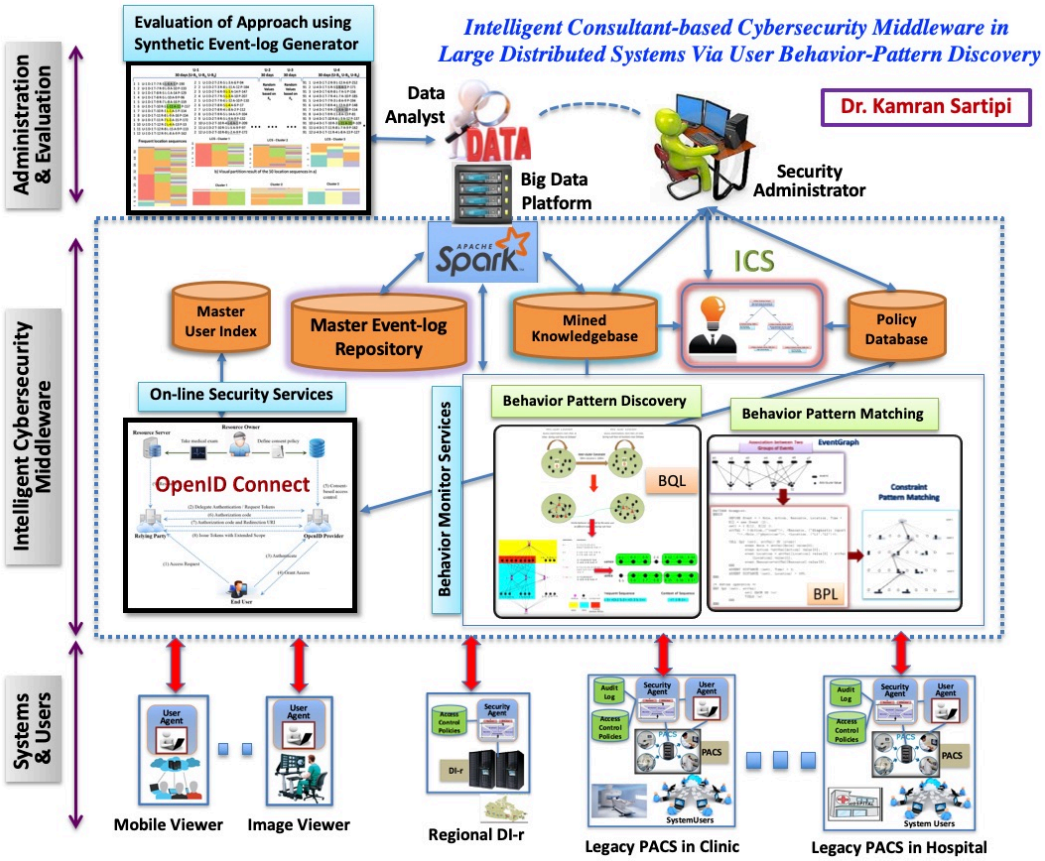
Kamran Sartipi
BSc, MSc (ECE), M.Math, PhD (CS), P.Eng

RESEARCH in CYBERSECURITY and BIG DATA ANALYTICS

Overall, my research activities span different aspects of provisioning an intelligent middleware infrastructure for knowledge-driven and customizable decision support systems through cloud and mobile services. The characteristics of this infrastructure are as follows:

- Utilizes different machine learning and data mining techniques to process large and heterogeneous datasets using Apache Spark platform to provide fast in-memory and real-time data analytics power.
- Contains a knowledgebase with the representation format and annotations appropriate to the target application domain.
- Utilizes virtualized intelligent decision services with the capability of exploring the knowledgebase in order to provide selective and non-overwhelming consultation guides to the users.
- Provides customizable generic agents in mobile devices which invoke cloud-based decision services to effectively assist the users and utilize visually encoded technologies as well as augmented reality.

I have applied this infrastructure to assist security administrators of distributed systems to identify malicious user-behaviors to enhance the security policies, which is described below.



Intelligent Consultant-based Cybersecurity Middleware Based-on User-Behavior Pattern Discovery

This research project (shown in figure above) provides a security middleware infrastructure for a large distributed medical system including several legacy and new PACS (picture archiving and communication systems) that desire to securely integrate their services and share medical images and reports. In such an environment, the authenticated and authorized users (trusted users) in one PCAS system can access to the resources of other systems with no proper security control that cause serious damage to such a sensitive information system. The security administrators cannot effectively monitor and control such situations due to huge volume of dynamically changing user behaviors. The objective of this research is to assist security administrators to enhance the security policies of such a large distributed system. To achieve this goal, a harmonized set of scientific techniques from data mining, knowledge engineering, constraint pattern matching, customizable agents, and decision support systems provide an effective and non-overwhelming consultation service for monitoring user behaviors and enhancing security policy rules. This service provides step by step consultation to the security administrator on the suspicious behaviors to identify affected users, resources, and locations precisely. A data analyst utilizes our behavior query language (BQL) to specify high-level clusters of events and their constrained relationships which allows for a focused and goal-oriented extracting user behavior patterns using machine learning techniques to populate a knowledgeable. Our behavior pattern language (BPL) is used to compose suspicious complex user-behavior patterns to be searched for approximate matches in the event-log repository to identify anomaly behaviors and the involved resources and locations. The proposed infrastructure also utilizes cloud-based open access authentication and authorization mechanisms for secure sharing of documents and resources among heterogeneous legacy PACS systems. Augmented reality mechanism is employed to provide intuitive, comprehensive and fast visualization aid to the security administer to navigate the generated patterns of behaviors.

Below, the sub-projects related to cybersecurity are briefly described with links to the PDF publications.

- **Intelligent middleware security provisioning.** A secure, central and service-based "intelligent middleware" consisting of: multi-agent technology (smart local agents and administrative middleware agents) for two-level decision-making process; central policy repository and management; central metadata repository for images; and a centralized authentication and decentralized authorization model **[**j12**,** j13**,** js4, c50, c48**,** c46**,** c45, c44, c42, c22**].**

- **Smart decision support systems.** This project aims at providing a new generation of decision support systems where mined-knowledge at decision points (as reminders, alerts, recommendations) will assist the physicians (for patient diagnosis) or administrators (for resource allocation) to make effective decisions. In this context, mined-knowledge refers to the extracted patterns and trends from clinical data using data mining techniques. This research covers both rule-based and flow-based decision support techniques [j10, j4, ch3, ch2, c43, c41, c31, c13, c11].

- **Knowledge-driven user behavior-pattern discovery.** This research provides a new generation of intelligent decision support systems that effectively assist the system administrators to obtain deep insight into the system user's dynamic behavior patterns in order to refine the existing security policies using a novel behavior pattern query language (BPQL) **[**j9**,** j14**,** j8**,** c49**].**

- **User behavior simulation environment.** An event-log generator engine is developed which receives administrator-defined user-behavior patterns using a pattern language and produces corresponding events in the context of noise events which allows us to effectively test and fine-tune the above techniques before applying them on the production event-logs. Currently we are experimenting with the audit logs from a distributed medical imaging system running at Mohawk College **[**j11**,** j14**].**

- **Enhancing data privacy in service oriented architecture (SOA).** This research enhances data privacy and security, reduces network traffic, and provides new enterprise level features. It introduces two new concepts "task service" and " service representative" in the SOA environment. Task service is a multi-component (model, knowledge, data) web service that can process the client data locally at the client side. Service representative is a generic agent at the client side that will be customized by the knowledge component and will execute the model component on both client and task service data. This approach will enhance the SOA architecture in different ways [j6, js2, js1, c39, c38, c36, c35, c34, c32].

- **Mobile eHealth.** This project provides efficient techniques to integrate different devices such as cell phones, tablets, and specialized devices such as pacemakers to be used seamlessly with other software services of the electronic health record (EHR) systems. In this approach data will be collected, maintained, analyzed and communicated using HL7 information models and messaging. [c35, c34].

- **Dynamic Analysis**. This research identifies the implementation of specific software functionality within a software system without any prior knowledge about the source code. The approach consists of applying specific sets of scenarios on an instrumented software system to extract execution traces. Next, sequential pattern mining algorithm and concept lattice analysis are applied to extract execution patterns and locate the target source code. We expanded this approach by applying it on service-oriented architecture (SOA) to measure the quality of web services in service selection and composition. [j5, c37, c26, c23, c20, c18, c17, c16, c15, c14, c12, c27].

- **Static Analysis**. This research addresses the design and development of an incremental software architecture recovery and evaluation environment using data mining techniques. The environment is interactive and provides: pattern-based architectural recovery using a query language and approximate graph pattern matching; optimization clustering; partitioning; and view-based architectural design evaluation. These techniques have been implemented within my Alborz toolkit, which are mentioned below [b1, ch1, j3, j1, c24, c10, c9, c8, c7, c6, c5, c4, c3, c2, c1].