Computer Vision

For additional info contact: ashour@il.ibm.com

Position:
Researcher in the Video and GIS Analytics group.

Project Description:
The Video and GIS Analytics group – part of the Multimedia Analytics department at IBM Research – specializes in advanced real-time video technologies. This unique group focuses on novel approaches for information overload and technologies in the areas of real-time rich-content streaming and management, video communication frameworks and geo-spatial situational awareness. The group develops a scalable framework for real-time connectivity, as well as a platform for off-line rich-media tagging search & retrieval of archived rich media assets.

In the area of video analytics, our group conducts research and develops novel computer vision algorithms (also using machine learning tools) for various problems such as scene text detection and recognition in natural videos and images, video segmentation, visual recognition and scene understanding.

The goal of the project is to develop and implement novel algorithms for solving various computer vision problems related to video (such as the ones mentioned above), with a special emphasis on robustness and efficiency.

We Offer:
An opportunity to work and conduct high-quality research with a leading research group in the area of video analytics and computer vision, on a topic that is on the cutting edge of research and technology.

Required Skills:

Graduate student (preferable) with good research and self-learning skills, as well as with some background in image and video processing, and preferably also in computer vision and machine learning. Good programming skills in MATLAB and C++ are also required (prior acquaintance with Open CV is a plus).

Full/Part time Position: Summer internship.

Summer Internship

For additional info contact: idanb@il.ibm.com

Position:
Researcher at the Smart Client Platforms group, in the Mobile Enterprise Platforms domain.

Project Description:
Enterprise mobility is about mobilizing the business organization. It involves the provisioning of tools, resources, and processes by the organization to allow its employees to effectively perform their tasks while on the road or away from their workstations. Typically, this involves an extension of business applications/solutions through the use of portable and mobile devices.

Mobile Enterprise Platforms deal with the technical capability to create, deploy and manage suite of enterprise apps to multiple, heterogeneous devices (iOS, Android, RIM) that connect securely to enterprise backend servers.

The Smart Client Platforms group specializes in the study and development of tools and platforms for developing, deploying, managing and analyzing enterprise applications using cutting edge Web 2.0, mobile and cloud technologies as well as advanced code analysis algorithms. The platforms we develop target skilled developers as well as knowledge workers and highlight operation simplicity, consumability, scalability, excellent user experience, while supporting high end security standards.

The intern will be exposed to frontend and backend technologies such as Ajax, REST APIs, JSON, and more.

We offer:
An opportunity to work and conduct high-quality research with a leading research group in the area of Mobile Enterprise Platforms on a topic that is on the cutting edge of research and the technology arena.

Required Skills:
Good research and self-learning skills, excellent programming skills including knowledge of Mobile Platforms, Web technologies.

Full/Part time Position: Summer internship.

Expressive speech: parameterization & transformation

For additional info contact: slava@il.ibm.com

General Background:
One of the key challenges in human computer voice interaction systems is the ability to manipulate synthesized speech, so that it can convey desired expressiveness and affect, while preserving high speech quality. The affect in speech is expressed by its prosodic characteristics (pitch, duration & energy) as well as by other speech manipulations (e.g. spectral envelope modification/ breathiness modification, etc.). Existing prosody modification techniques usually makes a simplistic assumption of spectral envelope preservation when modifying the prosody, while emotion modification is mostly based on prosody modification and to a lesser extent on some rule based spectral modifications.

Project Description:
We would like to explore data-driven approaches for prosody and emotion modification of speech (in parametric domain), provided a labeled and clustered emotional speech data corpus. We will start from prosody and emotion manipulation on natural prerecorded speech (e.g., part of the given labeled single speaker data corpus), and later move to synthesized speech manipulation.

We offer:
We offer an opportunity to work on an exploratory research project possibly leading to a scientific publication. Large data corpus will be provided for the research.

Skills:

- PhD candidate from EE or CS (or strong MSc student in advanced stages of his Master).
- Knowledge and Research Experience in Speech/Signal processing.
- Background in Speech Synthesis is an advantage.

Full/Part time Position: Full time student position for the summer period in the Haifa area.

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

Expressive speech: Multi-Modal Biometric Authentication

For additional info contact: Hagaia@il.ibm.com

General Background:
Multi-Modal Biometric authentication is the verification of a user's identity by means of a its physical traits or behavioral characteristics that cannot be changed easily, such as a voice, face, handwriting and fingerprints. Mutli-Modal Biometrics enables secure, robust and convenient authentication which is not available using only a single modality.

Project Description:
We would like to explore innovative approaches to improve the robustness and accuracy of our biometric engines and detection of spoofing attacks (user liveness detection). One such approach is the usage of Deep Neural Networks and an additional possible direction is to explore synchronous analysis of voice and the video live input in order to improve accuracy, robustness and liveness detection.

We offer:
We offer the opportunity to work on an exploratory research project possibly leading to a scientific publication in addition to contributing to the development of a novel mobile authentication solution.

Skills:

- PhD candidate from EE or CS (or strong MSc student in advanced stages of his Master).
- Knowledge and Research Experience in machine learning and either computer vision or speech/signal processing.
- Background in biometric identification or authentication is an advantage.

Full/Part time Position: Full time student position for the summer period in the Tel-Aviv area (preferred) or Haifa area (possible).

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

Expressive speech: Speech based Emotion Detection

For additional info contact: aharonsa@il.ibm.com

General Background:
Affective Computing is the study and development of systems and devices that can recognize, interpret, process, and simulate human affects (emotions, mood and feelings). Detecting and monitoring emotions in speech using its non-verbal content is a key technology in affective computing and affect aware conversational systems.

Project Description:
Our goal is to go beyond the recognition of basic emotions (anger, sadness, happiness etc.) and develop algorithms for detecting complex/secondary emotions from speech signals (e.g. disappointment, frustration, satisfaction), which are more typical in human machine dialogue.

In addition to conducting algorithmic research the work may include running lab experiments and collection of data for the research.

We offer:
We offer the opportunity to work on an exploratory research project possibly leading to a scientific publication.

Skills:

- PhD candidate from EE or CS (or strong MSc student in advanced stages of his Master).
- Knowledge and Research Experience in speech/signal processing and/or machine learning.
- Experience in software development in Windows environment.

Full/Part time Position: Full time student position for the summer period in the Haifa area.

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

Summer Internship

For additional info contact: dannygu@il.ibm.com

General Background:
In 2006 a team of researchers in IBM took upon themselves the challenge of developing a fully automatic system that will compete in the popular quiz show Jeopardy! It took them over 5 years and in 2011 the system, which was named Watson, participated in a live Jeopardy! Show competing against the all times champions. Watson won the game while demonstrating unprecedented text analytics capabilities. Since this event, IBM has been building on this technology to open new business directions for the company. At the same time researchers at IBM asked themselves what would be the next challenges for Watson.

Project description:
While Watson demonstrated impressive text analytics, it is still confined to answer factual questions where typically there is a right or wrong answer. However, most of the questions that we ask in life are more complex and are influenced by biases and different points if view. For example: "Should we ban smoking?", "Should I rent an apartment or buy one?", "Should IBM make a partnership with Apple?". IBM Debating Technologies is a project which aims to address such scenarios. At a very high level, we are developing a system and a set of tools, which will assist humans in situations where debate and reasoning is required. The system, given a topic under dispute, generates arguments which either support or contest the topic. A demonstration of initial capabilities can be seen here (starting around minute 45):
https://www.youtube.com/watch?feature=player_embedded&v=6fJOtAzICzw

We offer:
An opportunity to be part of an interdisciplinary, global team, working on a cutting edge technology which is highly exploratory.

Skills:
- Proven background in natural language processing and/or machine learning
- Programming experience in Java is an advantage.

Full/Part time Position: Summer internship.

Summer intern in the field of Medical Imaging

For additional info contact: flora@il.ibm.com

Position:
Computer Vision student at the Medical Imaging Analytics Group.

Project Description:
The Medical Imaging Analytics Group– part of the Multimedia Analytics department at IBM Research – specializes in advanced image processing technologies. The project scope includes invention and design of new computer vision methods in the field of medical imaging.

Position requirement:
We are looking for researchers in the field of computer vision and machine learning. The position will involve analysis of radiological images and clinical data.

We offer:
An opportunity to collaborate and publish with top researchers at IBM Research in the area of medical image processing, on a topic that is on the cutting edge of research and technology.

Required Skills:
The position is intended for MSc or PhD students in Computer Science/Electrical Engineering/ Mathematics with focus on computer vision, machine learning or image-processing.
Research capabilities, with strong theory/algorithm background and very good understanding on how to apply advanced knowledge to solve real problems.
Fluent in MATALB.
Industry experience is an advantage.

Full/Part time Position: Summer internship.

Proactive Location Intelligence

For additional info contact: bnayahu@il.ibm.com

Position:
Summer internship at the Smart Decision Solutions Group.

Project Description:
With the pervasive adoption of location aware technologies and other sensors in mobile devices, connected cars, environmental sensors, and other software and hardware architecture, today's smarter systems need to be able to sense, analyze, monitor, predict, and response to space- and time- based situations. Our group aims to advance the state of the art in development of comprehensive set of technologies that facilitate building smart systems than can react, predict and take proactive actions to space-time situations. We apply our technologies for solving real-world client challenges across a wide range of industries such as Travel & Transportation, Logistics, Maritime, Insurance, Natural Resources and Retail. In addition we apply our technologies to address societal challenges in areas such as Smarter Cities, Urban Planning, Urban Mobility, Multi-modal Transportation, Environment, Open Data, and Citizen Engagement. We work in close partnership with other business units in IBM to deliver a standards based technological support for the IBM software platform, solution and services businesses.

References: https://www.research.ibm.com/haifa/dept/services/sds.shtml

We Offer:
An opportunity to join R&D activities in the world's largest IT research organization.

Required Skills:
Good research and self-learning skills, programming skills (primarily in Java) and a passion for innovation that matters.

Full/Part time Position: Summer internship.

Analysis of Pareto Frontiers

For additional info contact: amirka@il.ibm.com

General Background:
Pareto frontiers are an important class of multivariate datasets. Consider almost any decision-making problem: it can be modeled as a set of points, where each point represents a possible alternative for the decision maker to select from. Moreover, each alternative has its own merits and weaknesses; this is modeled by giving values for the alternative in several pre-determined criteria, so that each point is n-dimensional. All in all, the decision problem is modeled by a set of n-dimensional points. The Pareto frontier is a subset containing all those points $\vec{p}$ that are maximal, or locally optimal – i.e., there is no other point that is greater-or-equal than $\vec{p}$ in all coordinates (given that greater values are better). On the basis of those specific criteria, there is no reason for the decision maker to select a point that is not on the frontier. However, mathematically, all the alternatives on the frontier are equally attractive.

Position:
Student Researcher.

Project Description:
We believe that in some cases, some of the points on the frontier can be understood to be more attractive than others. The aim of this project is to develop a rigorous mathematical model and devise mathematical and algorithmic methods to assess which are the more attractive points on the frontier. The project also requires implementing these ideas in a high-level programming language such as Python.

We offer:
We offer an exploratory research project possibly leading to a scientific publication, in addition to participating in developing a novel solution with a clear productization roadmap.

Skills:
- Strong analytical thinking.
- Strong background in mathematics / mathematical sciences.
- Experience in mathematical (esp., probabilistic) modeling is an advantage.

Full/Part time Position: Full-time student position for the summer period.

Mining User Interactions

For additional info contact: amirka@il.ibm.com

General Background:
From the interactions of users with an application or website one can learn certain things about the user, e.g., his goals and needs, and about the application's quality and characteristics. Moreover, in the context of applications that handle data, one can sometimes implicitly infer certain things about the underlying dataset – solely from the way that users manipulate it.

Position:
Student Researcher.

Project Description:
This project aims at devising ways for making valuable insights from the interactions of users with a certain type of applications (the exact nature of these applications, and the insights that we expect to extract cannot be announced at this stage). The project includes determining exactly which interaction features to collect for the problem at hand, and applying machine learning algorithms and data mining techniques for the extraction of valuable insight from the collected data. The project also requires implementing these methods in a high-level programming language such as Python or R.

We offer:
We offer an exploratory research project possibly leading to a scientific publication, in addition to participating in developing a novel solution with a clear productization roadmap.

Skills:
- Strong analytical thinking.
- Strong background in data mining and machine learning.
- Experience in data analysis is an advantage.

Full/Part time Position: Full-time student position for the summer period.

Summer Internships

For additional info contact: liorli@il.ibm.com

Company: IBM - HRL

Position:
Stream pattern recognition over continuous sensor data.

For graduate students / undergraduate students:
Graduate students.

Job description:
The proposed research project will focus on the adaptation of on-line machine learning techniques for personalized pattern recognition over continuous data streams, with each pattern being a series of data which can be correlated with a previously observed sequence of data. For example, over excessive effort prevention may require constant learning of person's activities to minimize false positive detection while adapting to the circumstances in which the person is at. Similarly, the ability to train the system to enable interactions via hand gestures (e.g., for field operators) may require a convenient way to add and adopt the system ability to recognize various gestures being of interest to the user.

Requirements:
The project may be most appealing to students who are interested or who are already engaged in pursuing research that employs machine-learning/data/process-mining techniques for the purpose of on-line stream processing. The intended work in this project will start with a survey of existing techniques to be employed for the aforementioned use cases, followed by a proof-of-concept implementation that compares between the ones identified as the most prominent. The scientific work will be supervised by lead scientists in HRL, and will be planned with an aim to publish an article in a top venue.

Full/part time position: A full term summer internship.

Summer Internships

For additional info contact: liorli@il.ibm.com

Company: IBM - HRL

Position:
Adaptive sampling for activity recognition over continues sensor data.

For graduate students / undergraduate students: Graduate students.

Job description:
Most of nowadays techniques for activity recognition are based on the partitioning of sensor data (e.g., accelerometer) into a sequence of sampling windows. Each such window is then extracted a set of features that are used as the input for the employment of various machine learning techniques for the purpose of activity recognition. Such techniques are used to develop the models that are then used to classify the various activities. As a result of this methodology, the partitioning into windows may be crucial to the entailed accuracy of these models. The focus in this work will be on the development of a proof-of-concept which employs an adaptive technique that dynamically adjusts the sampling rate to math with the actual activity at hand in order to improve the accuracy of activity recognition.

Requirements:
The project may be most appealing to students who are interested or who are already engaged in pursuing research that employs machine-learning/data/process-mining techniques for the purpose of on-line stream processing. The proposed research project will survey existing techniques for activity recognition, with a focus on their underlying technique for window partitioning. Particularly, it will explore algorithmic approaches to the determination of such windows in the context of on-line data streams, and the accuracy effects that may be attained from the adaptation of such windows to the input data. The scientific work will be directed by lead scientists in HRL and will be planned with an aim to publish an article in a top venue.

Full/part time position: A full term summer internship.

Parser for Advanced Processor Verification Template Language

For additional info contact: marcus@il.ibm.com

Company: IBM Research Labs – Haifa.

General background:
Random-based stimuli generators are widely used in the hardware industry for the verification of hardware components such as processors. The input to a processor-level stimuli generator is a test template, which describes at a high level the desired characteristics of the test cases to be generated. Given a test template as input, the stimuli generator produces a large set of architecturally valid tests that satisfy the template request by randomly filling in all unspecified details. These test-cases are then executed on the design under test to verify that it works as expected.

At the heart of this approach lies a sophisticated template language for defining test templates. This language must be sufficiently expressive to allow users to specify details of what they would like to see in the generated tests on the one hand, as well as allow users to omit all irrelevant details that they would like to be filled in by the stimuli generator. This second requirement is what makes template languages more complex than standard programming languages and demands new algorithms to ensure correct parsing and handling of templates.

Project description:
In this project we will explore and implement advanced parsing techniques which can handle the complexities introduced by template languages. We will investigate using constraint satisfaction (CSP) for disambiguating various language constructs, as well as sophisticated graph-based approaches to handle language inter-dependencies. Successful results of this research may be submitted for publication.

Required skills:

- Good research and self-learning skills.
- Good programming skills (preferably C++).
- Graduate student is preferred. Outstanding undergraduate can also apply.

[1] Katz, Y., Rimon, M., Ziv, A.: Generating instruction streams using abstract CSP. Proceedings of the 2012 Design, Automation and Test in Europe Conference. pp. 15-20 (March 2012).
[2] Katz, Y., Rimon, M., Ziv, A.: A novel approach for implementing micro-architectural verification plans in processor designs. Proceedings of the 8th Haifa Verification Conference. pp. 148-161 (2013).

Full/part time position: Full time summer internship.

Social Engagement Analysis Research in Social Media Technologies Group

For additional info contact: Shiri Kremer-Davidson, Social Technologies group. shiri@il.ibm.com,
http://researcher.watson.ibm.com/researcher/view.php?person=il-SHIRI
http://www.research.ibm.com/haifa/dept/imt/ct_st.shtml

Position:
Researcher at the Social Technologies group - IBM Haifa Research Lab.

General background:
Social networks such as Twitter, Facebook, Instagram etc., have become very dominant communication tools among people. Similarly to their counterparts on the Web, social media applications have also emerged inside organizations. Yammer, Chatter, and IBM Connections are a few prominent examples. Through functionality like personal walls, forums, wikis and blogs, they provide a rich engagement platform through which employees can interact, share their experience and knowledge, learn new insights and widen their acquaintance network. Moreover, it opens an additional window of opportunity for employees to get a voice within the organization.

Recent studies prove the importance for companies to have their employees socially engaged.

Internship proposal:
In this project, the goal is to develop *engagement analytics* over the enterprise social network activity graph. Such analytics can include identifying employees' social behavior patterns, providing recommendations to employees on how to engage in a more efficient manner, identifying insights to management from the company social network, etc. Our goal would be to both to develop a working prototype and to publish a scientific paper.

Requirements:
Graduate students with strong engineering skills and excellent research skills who can work as part of a team. Java programming knowledge is a must, due to the short term of the project. In-depth knowledge of statistical analytics and/or machine learning is an advantage.

Full/part time position: The project fits a 3-month internship on a full position basis.

Graph Partitioning

For additional info contact: Roy Levin, Ph.D., Social Technologies group, royl@il.ibm.com

Haggai Roitman, Information Retrieval group, haggai@il.ibm.com

Position:
Researcher at the Social Technologies group - IBM Haifa Research Lab.

General background:
In recent years graph databases are drawing more and more attention from both industry and academia mostly due to the proliferation of Online Social Networks (OSNs) and Linked Data. Due to the sheer volume of data, in many examples, it is no longer possible to effectively store the entire graph on a single machine. As a result the graph needs to be partitioned across multiple machines. Hence the issue of balancing the graph partitioning in a streaming setting is a key problem to enable scalable and efficient computations. Yet many graph database systems provide only basic partitioning strategies such as randomized partitioning or partitioning based on transactional affinity. More advanced partitioning techniques are based on a heuristics which either places the newly arrived vertex in the cluster with the largest number of neighbors or alternatively, in the cluster with the least number of non-neighbors or some combination thereof. These methods provide suboptimal partitioning resulting in poor query processing as each query may issue traversals that require data residing on multiple machines or unevenly balancing the workload among the different machines.

Project Description:
In this project we will implement, using an open source graph database called Titan, a method which performs live migration of vertices between machines based on actual traversal queries issued to the system. The goal will be to show that our proposed method reduces query processing time when compared to other known techniques for partitioning such as those mentioned above.

References:
FENNEL: Streaming Graph Partitioning for Massive Scale Graphs - WSDM 2014.

We offer:
An opportunity to work and conduct high-quality research in one of the world's leading research groups in the area of Social technologies and Information Retrieval on a topic that is in the bleeding edge of research and technology.

Required Skills:
Graduate students with strong engineering skills and excellent research skills who can work as part of a team. Java programming knowledge is a must, due to the short term of the project. Web development skills are an advantage and so is existing publication experience.

Full/part time position: Summer internship **-** 3-month internship on a full position basis.

Real-Time Trend Detection in Social Networks

For additional info contact: Roy Levin, Ph.D., Social Technologies group, royl@il.ibm.com

Position:
Researcher at the Social Technologies group - IBM Haifa Research Lab.

General background:
Detecting trending events in online social media streams has become a very popular topic in recent years. However, since the same set of trends may not interest all users to the same degree, solutions have also begun to emerge for the problem of promptly recommending trending entities within some predefined personalized context. However, the problem of finding trends related to a search query which is NOT known in advance has not yet been studied. Such a setting would allow users to promptly get relevant trends as they issue a new search query over a dataset of documents. For example, recommending trending applications when a user searches for ``children games''. This presents a new challenge as, unlike existing methods, there is no predefined context for which relevant trends can be calculated in advance.

Project Description:
We have designed the outline of an algorithm for detecting such trends, in real time, over streaming social media data within the context of a search query. For our purposes we define an entity as trending if it is expected that it will be more frequently mentioned within the near future, indicating that it is becoming more important. Hence, using real data from an online social network, we wish to evaluate the algorithm based on its ability to predict an actual increase in the level of activity within entities that are found to be trending in the given search context.

We offer:
An opportunity to work and conduct high-quality research in one of the world's leading research groups in the area of Social technologies and Information Retrieval on a topic that is in the bleeding edge of research and technology.

Required Skills:
Graduate students with strong engineering skills and excellent research skills who can work as part of a team. Java programming knowledge is a must, due to the short term of the project. Web development skills are an advantage and so is existing publication experience.

Full/Part time Position: Summer internship - 3-month internship on a full position basis.

Social Analytics over email and social media

For additional info contact: Inbal Ronen, Social Technologies group, inbal@il.ibm.com

Position:
Researcher at the Social Technologies group - IBM Haifa Research Lab.

General background:
Social media sites such as Twitter and Facebook, are becoming more and more dominant on the web. Users share their activities, opinions, photos, and videos, and in parallel get similar updates from their social environment. Moreover web mail sites such as Google and Yahoo add more and more social features to their sites. Similarly to their counterparts on the Web, social media applications have also emerged inside organizations. Yammer, Chatter, and IBM Connections are a few prominent examples.

Project Description:
In this project, the goal is to develop social analytics over the corporate's mail or social media sites. Such analytics can include prioritizing and categorization of a user's mail, identifying to-do's in mails, recommending related context and people to the user's current work context and more. Machine learning and Information retrieval would be among the core algorithms in this project. Our aim is to both develop a working prototype and publish a scientific paper.

We offer:
An opportunity to work and conduct high-quality research in one of the world's leading research groups in the area of Social technologies and Information Retrieval on a topic that is in the bleeding edge of research and technology.

Required Skills:
Graduate students with strong engineering skills and excellent research skills who can work as part of a team. Java programming knowledge is a must, due to the short term of the project. Web development skills are an advantage and so is existing publication experience.

Full/Part time Position: Summer internship - 3-month internship on a full position basis.

I/Q generator at 8-15 GHz

For additional info contact: Danny Elad (dannye@il.ibm.com) and Oded Katz (katzo@il.ibm.com)

General Background:
Our team designs and develops silicon based transceivers at E-band and V-band frequency range for communication applications. A single silicon based transceiver chip offers high integrability and low cost products. The I/Q generator is a basic building block for such transceivers used to generate high order quadrature amplitude modulations.

Project Description:
The intern will be responsible of designing an I/Q generator at 8-15 GHz frequency range. The technology used in the project will be SiGe8HP (Silicon –Germanium Heterojunction Bipolar transistors). The project includes all design stages, starting from schematic, physical design, and implementation for testing.

We offer:
The project demands high level of analog and RFIC circuit design. The versatility of the circuit may require innovation, while physical implementation might lead to a scientific publication.

Skills:

- Background in analog circuit design or RFIC design.
- Experience in analog circuit design environment.

Full/Part time Position: Full time student position for the summer period.

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

Security and Privacy Services for Bluemix

For additional info contact: borisr@il.ibm.com

General Background:
IBM Bluemix is a cloud platform as a service (PaaS) developed by IBM. It supports several programming languages and services as well as integrated DevOps to build, run, deploy and manage applications on the cloud.

Project Description:
The project goal is to learn Bluemix environment and understand how to extend it with Security/Privacy related technologies developed at HRL. The anticipated outcome of the project would be:

- Adding one of the security/privacy related technologies developed at HRL as a service in Bluemix.
- Report with guidance/best practices for leveraging Bluemix for developing secure applications.

We offer:
We offer exploratory Research project leading to developing a novel solution with a clear productization roadmap.

Skills:

- Good research and self-learning skills.
- Strong background in Java, Unix/Linux.

Full/Part time Position: Full time student position for the summer period.

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

BigData security analytics of Cloud Systems in the cloud

For additional info contact: shelly@il.ibm.com, yaron@il.ibm.com

General Background:

Any computer system, and cloud services in particular, produces logs which are a great source of understanding the system: who did what? when? where? and how?
BigData real-time analysis of these logs is important in order to audit the system for security, monitoring and other purposes. Moreover, logs of cloud systems are considered today a canonical example for BigData, hence analyzing this data requires highly scalable storage and analytics infrastructure.

Project Description:

The project purpose is to perform an advanced analytics of logs produced by various cloud services using large scale cluster computing (such as Apache Spark and Hadoop), and develop a new solution that will be combined in the cloud to audit the logs in order to enhance the cloud security. The project will focus on OpenStack, which is today's most popular cloud open source project.

We offer:

An opportunity to conduct high-quality research possibly leading to a scientific publication in addition to participating in developing a new solution. We also offer an opportunity to explore with state-of-the-art scalable cloud systems as well as cloud analytics tools.

Skills:

- Strong algorithmic background.
- Programming skills (in Python, Scala or Java).
- Familiarity with scalable clustered systems.
- Experience in machine learning and data mining is an advantage.

ROBUST NON-FACTOID QUESTION-ANSWERING

For additional info contact: oferl@il.ibm.com

In recent years considerable attention has been given to the problem of question answering. However, most of the work has focused on automatic question answering frameworks for factoid questions rather than non-factoid questions. This task of automatically answering non-factoid questions is still open for deep investigation.
In this project we develop a robust and scalable methodology for non-factoid QA using machine learning and natural language processing techniques which are applied to big data sets to improve search, ranking, question-answering and many other tasks.

We are looking for talented and motivated graduate students who will examine new directions for this project, specifically in the domain of word embedding and neural networks. One possible direction is to enrich the data set by creating paraphrases of the questions.

Join in -

- If you are a computer science or engineering PhD/MSc student.
- If you master programming languages – Matlab / R / Python / Java
- If you are creative.
- And ready for hard work in a pleasant environment :-)

Some more details:

- 3-4 months of a PhD / MSc student during the summer.
- Computer Science / Engineering students.
- Proficiency in machine learning and data analytics.
- Background in Natural Language Processing - an advantage.
- Project involves research and development of methods in cognitive computing and machine learning and applying them to real world data.
- An option for submitting the work as a paper in the weeks following the internship.
- Get to know IBM's recent tools such as BlueMix.

Bringing Big Data, Cloud, Mobile and Human Interaction into an Holistic Cognitive Solution

For additional info contact: oferl@il.ibm.com

Background:
Big Data, Cloud, Mobile and Human-Computer Interaction (HCI) are all important ingredients in the formation of an end-to-end machine learning and cognitive solutions.
However, creating a solution that consists all of these components requires variety of skills and proficiencies.

Project Goal:
In this project, we will construct a re-usable modular framework that allow a fast development and deployment of end-to-end cognitive solutions by minimizing the overhead associated with selection of technologies, learning curve, integration and deployment.

Method:
We will first review and analyze alternative technologies and tools for the different framework tasks (Server Side, Client Side, Machine Learning, Visualization, Cloud, etc...) - 4 weeks
Then we select the appropriate components and validate them by creating a sample end-to-end cognitive application - 4 weeks
In the next step, we will convert the application into a modular template by decoupling and isolating the application logic from the framework - 2 Weeks
And document our work by writing a step-by-step guide on how to use and configure the framework for fast creation and deployment of future cognitive solutions - 2 Weeks

Looking for...
We are looking for a talented and motivated graduate students who will take the major role in the execution of this project,
The candidate does not need to be an expert in any of the above mentioned technologies, Nevertheless a background in development of both client and server side is required and a broad perspective of modern technologies and software development environments is a big advantage.

Join in:
- 3-4 months of a PhD / MSc student during the summer
- Computer Science / Engineering students
- At least two year of past software development experience out of the academy.
- Good Java programming capabilities.
- Master both client side and server side software development.
- Good problem solving and self learning capabilities.
- Broad perspective of modern technologies and software tools - Big advantage
- During the project, you will get to know IBM's recent tools such as SPSS Modeler, Bluemix, etc...

- Possibility to work either in Haifa or in Tel-Aviv.
- Ready for hard work in a pleasant environment :-)

Research and Development in Constraint Programming

For additional info contact: Yael Ben-Haim yaelbh@il.ibm.com

Company: IBM Research Labs – Haifa.

General background:
The Constraint Satisfaction team focuses on research and development in the area of constraint satisfaction problems (CSPs). Our main asset is the IBM Constraint Solver, a robust, general-purpose, state-of-the-art tool that has been used for more than a decade in modeling and solving many complex constraint problems. Our department has long-standing expertise in CSP algorithms and modeling. Our aim is to provide value to IBM through the application of constraint solving to various domains along with close interaction with the academic community. We work closely with our IBM partners to find the best constraint or optimization model for their domain and to apply the best heuristics when solving this model.

More details about our activities can be found on our website:
https://www.research.ibm.com/haifa/dept/vst/csp.shtml

Proposal for internship project:
The goal of the project is to research in the area of constraint satisfaction problems, by enhancing our solver with new capabilities. Examples include:

- Developing novel algorithms for solving constraint problems.
- Experimenting with various types of solving heuristics.
- Exploring algorithms to find unsatisfiable cores.
- Exploring invocation of a SAT solver from the CSP solver.

Project deliverable / Publications:
A successful realization of the project will produce a C++ implementation of the new algorithms, integrated with our solver. In addition, the project may produce a paper that will describe the conducted research. The paper will be submitted to one of the top conferences in the area of Constraint Programming.

Required skills:
- Good research and self-learning skills.
- Good programming skills (preferably C++).
- Graduate student is preferred. Outstanding undergraduate can also apply.

Full/Part time Position: Full time summer internship.

Learning Process Structures for Security Oriented Anomaly Detection

For additional info contact: adir@il.ibm.com

General Background:
Cyber security is a field of ever growing demand and sophistication. Machine-learning and specifically anomaly detection techniques are commonly used tools in this domain. The typical behavior of the system is learned and deviations from this behavior (i.e. anomalies) can be highlighted as possible security breaches.
However, most current techniques focus on creating statistical models over a given set of features and the detected correlation of these features, or on modeling the flow of these features over time. We would like to focus on the *processes* that the changes in the features exhibit and then measure the deviation of an observed process from the modeled process.

Project Description:
We will analyze existing structured logs that list database accesses made by applications for the purpose of detecting anomalies in these logs that potentially indicate security breaches.
We will study an open tool (ProM) developed by a research team at the Eindhoven University of Technology that can create models of processes given logs that report the events related to the process. We will then use the process models to assess the probability and hence the anomaly of observed events.

We offer:
We offer exploratory Research project in collaboration with researchers at the IBM research Labs at Haifa and also at the Eindhoven University of Technology. The research will possibly lead to a scientific publication in addition to participating in developing a novel solution with a clear productization roadmap in the domains of both security and in Internet-Of-Things.

Skills:

- Background in Software-Engineering and Java.
- Basic background in data analytics / machine learning.

Full/Part time Position: Full time student position for the summer period.

Please apply on line- https://www.research.ibm.com/haifa/careers.shtml

Semantic Mediation

For additional info contact: Dr. Uri Shani, shani@il.ibm.com, 0546976282. For the SET group in HRL (Manager: Henry Broodney)

Problem domain:

Semantic Web proposes to assign semantics to resources on the internet whereby the resources are defined as graphs in RDF (Resource Description Framework) format, and semantics are defined as ontologies in OWL (Web Ontology Language) specifications, also using RDF format. OWL applies first order logic axioms so as to infer new facts from the given facts of an RDF dataset.

We in IBM have through our participation in several EU projects in complex model based systems engineering (MBSE), have developed a platform on which MBSE models for cyber-physical systems (e.g., aerospace and automotive products) use these semantic web approaches to enable sharing of models among different tools. That is a serious problem in this domain and it applicable to many other domains in modern IT where information must be shares across domains.

Semantic mediation is a simple principle where an RDF model can be mediated from one representation which complies with one ontology, to another RDF model that matches a different ontology. In the tool interoperability case, each tool can export and import its internal model representation to and from RDF, using semantic terms based on its own OWL ontology. The platform transforms that RDF to another RDF that is compliant with another tool working similarly to the first one. Thus, a model in one tool is shared with another tool.

The semantic mediator interprets rules written as OWL axioms, and on which it reasons to produce the mediated output model from the input one.

Work description:

The student will work with our researchers to study the rules specifications that are used by the IBM system, and help us to understand the theoretical foundation of this process, where it matches the standards of the OWL logic system, and where it may need more sound foundation to support it. As a practical system, we know that the pure OWL reasoning approach is inefficient, and we have developed a very efficient engine to do the job, and the student will help us to define the foundation of this practical approach and match it to the pure logic theory on which OWL is based.

Required Background:

The student is expected to have background as theoretical logician, where the study of OWL and RDF can be completed while on board. Yet, programming ability in Java by which implementation of ideas as well as experimenting with a working alternatives is a plus we will consider.