Förslag till insats inom Forskning för innovation 2016 (VP5)

Suggested Project within Research for Innovation 2016 (for Activity Plan 5)

Detta förslag berör/involverar följande forskningsmiljöer

This proposal concerns/involves the following research environments

[X] EIS   [ ] CIEL   [ ] CVHI   [ ] Annan / Other ----------------

Insatsnamn (Ange gärna även en akronym för insatsen) -- Project name (possibly including an acronym)

Self-Awareness and Self-Monitoring for Innovation (SASMI)

Subproject 1: Data-Driven District Heating Anomaly Detection (D^3HAD)

Subproject 2: Self-Organising Diagnostics in Industrial Networks (SODIN)

Subproject 3: Intelligent Toilet for Smart Home (iToilet)

Subproject 4: Fleet-based Vehicle Diagnostics (FVD)

Typ av insats -- Type of project

[ ] ProSpect   [ ] HÖG-project   [ ] HÖG-litet – HÖG-small

[X] Synergy   [ ] Profile   [ ] Profile+

[ ] Företagsforskarskola – Industrial graduate school   [ ] Strategic recruitment (professor, assist. prof.)

[ ] International guest professor   [ ] Adjunct professor/ adjunct assistant professor

[ ] Forskningsledare – Research leader (by invitation)   [ ] Adj. av FoU-chef/spec. – Adjunction of research director/specialist

[ ] Expert competence for innovation   [ ] Avans (development of education at advanced level)

[ ] IT i högre utbildning – IT in higher education   [ ] Annan -- Other ________________

Tilltänkt projektledare och deltagare (ange även om ny doktorand/anställning planeras inom insatsen) – Intended project leader and participants (also, specify if new PhD student or other new employment is planned within this project)

Project leader: Prof. Antanas Verikas

Sub-project leaders: Dr. Anita Sant'Anna (1) , Dr. Slawomir Nowaczyk (2&4), Dr. Stefan Byttner (3)

Participants: Thorsteinn Rögnvaldsson (Prof), Sven Werner (Prof), Hasan Nemati (PhD student), Martin Cooney (Dr), at least one newly employed PhD student and/or Postdoc

Ange medverkande företag, kontaktperson samt eventuella andra organisationer som kommer att medverka i insatsen (Observera att samproducenter ska vara vidtalade och ha accepterat att bidra till utveckling av insatsen. Notera också att KK-stiftelsens näringslivsdefinition ska följas, se tex anvisningar för HÖG14)

List participating companies and their respective contact persons, as well as other organisations that are intended as partners. (Please note that co-production partners must have been contacted and have accepted to contribute to the further development of the project plan. Please also note that, in order for a partner to count as a matching industry partner, the criteria used by KKS must apply. See information in, e.g., the HÖG 14 directives).

Öresundskraft – Henrik Gadd <Henrik.Gadd@hh.se>
HEM – Alexander Örning <Alexander.Orning@hem.se>

HMS – Michal Lysek <mly@hms.se> & Erik Halvordsson <erh@hms.se>
TBD (one of HMS customer companies, to be decided)

Villeroy & Boch Gustavsberg – Stefan Ljung & Charlotta Skoog
ACREO Swedish ICT
Other contacts are investigated (Microsoft Sweden, Bill and Melinda foundation), companies that can deliver sensors/embedded systems (ObserveMedical AB)

Volvo Group Trucks Technology – Rune Prytz <rune.prytz@volvo.com>

Beskriv hur förankringen med externa partner har skett
Describe how the preliminary support from external partners has been achieved

Based on preliminary discussions, the companies will contribute to the project in kind. Their contributions will cover meetings, requirements specification, formalization of domain knowledge, as well as prototype development, support with data collection, experiment setup and data analysis. Companies will also provide equipment necessary for conducting the research.

Preliminär frågeställning/problemområde som adresseras inom insatsen
Preliminary research question or problem area that is addressed in the project

We have a long term research cooperation with Volvo AB on self-monitoring systems, specifically buses and trucks. The goal is to enable a fleet of vehicles to build up a knowledge base about itself, in terms of on-board signal signatures and recorded maintenance operations, and how to use this to predict maintenance needs. The system should do this as autonomously as possibly, i.e. without requiring much interaction with human experts. The motivation is that future (and already today's) systems generate large amounts of (often streaming) data. This is a potential for improved analysis and operation of these systems but it is not a feasible (scalable) approach to require human experts to deal with the data, as current solutions do. The approach we have taken with the Volvo vehicles is to have software agents operate on-board, which hunt for "interesting" relationships or representations of vehicle operation. These representations are then transferred to a central server, where a self-organized algorithm looks for deviations. We have shown that this can be used to warn for several failing components on the vehicles, without the use of human expert knowledge. The challenges in the work have been how to handle human generated data (maintenance records) with their uncertainties, how to automatically find good representations of signal streams and how to compare these to each other.

This Synergy project aims to extend this approach to other domains: district heating, manufacturing and healthcare. We also intend to continue the research in the transport systems domain, and to use experiences from other industries to improve it further.

The operation of many modern systems, such as assembly lines in manufacturing factories, depends to a very high degree on the industrial networks connecting all the elements together and the data flowing in these networks. Those networks are typically thought of as passive parts of the infrastructure, but the data that flows through them can be used for much more. No two systems/setups/customers are exactly the same, but there are enough similarities that those data streams can be automatically analyzed and provide valuable information, e.g. for deviation and diagnostics purposes, taking into account a large number of intrinsic and extrinsic factors, as well as daily and seasonal variations.

District heating systems serve extremely large and complex networks, including many thousands of customers. The operation of such systems similarly depends on a large number of intrinsic and extrinsic factors, as well as daily and seasonal variations. Also here the system is further complicated by the fact that no two costumers are identical. The detection of faults and inappropriate configurations, which affect the efficiency of the system, is difficult and requires careful monitoring of individual customers and well as comparisons between similar clusters of customers.

Providing a toilet with a combination of an embedded system and appropriate sensors makes it possible to continuously make urine sample measurements. Through these measurements, a more or less advanced data analysis can be performed to monitor trends and detect correlations. It would for example be possible to follow the potassium or glucose level in individuals over time and changes or deviations from normal levels, either for the individual over time or for the individual versus a group of other individuals. A daily monitoring of urine status can lead to early detection of problems and thus also earlier dealings with the problems, especially if most likely cause for differences could be learned from documentation in medical health records.

Commercial vehicles such as trucks, buses or construction equipment are expected to operate for long periods of time and often work in remote areas, in addition to being mobile. There is a need for increasing uptime of such vehicles, while at the same time lowering development time and cost. This requires new, data-driven approaches to monitoring and fault detection. One example solution is a fleet-based, self-organizing diagnostics system that uses the data streams on the on-board CAN network and fleets of vehicles to learn the "normal" behavior.

In particular, we are concerned with the following research question, present in all four domains: how to utilize various available streams of signals (district heating, power meters, toilet measurements, CAN messages, etc.) to automatically define normality, detect deviations and learn signatures of faults and inappropriate operation, across multiple domains. This involves finding methods to autonomously rank data streams that are worth monitoring, generating and evaluating features describing those streams, methods for autonomous detection and quantification of "deviation", and methods for mining human curated text-based data (health records, maintenance histories, heating system inspections, etc.) and matching these to changes in the data streams.

Beskriv frågeställningens/problemområdets relevans för näringsliv respektive akademi
Describe the relevance of the research question / problem area for industry and for academia

This project addresses very similar problems in four different areas/industries: district heating, production/manufacturing, healthcare and transport systems. These are very distinct areas but with common questions that would profit from autonomous methods for monitoring and deviation detection in different ways. The term "autonomous" must be emphasized here: in a society with data streams everywhere it is impossible to have human experts overlook all these data streams and define how they should be represented. The monitoring/data mining systems must themselves be able to generate potentially interesting features (i.e. compressed representations of the streaming data) and over a long time evaluate whether these are useful or not. The domain we target share many important characteristics, leading us to believe that distributed deviation detection methods will be useful in all of them.

TRANSPORT SYSTEMS: Maintenance of commercial equipment is becoming more and more important as industries move towards just-in-time processes and the profit margins become slimmer. With the expectations of high uptime over long periods of time, it is useful to accumulate information over equipment lifetime. Such information, however, cannot be analysed by hand, and instead require data mining approaches that are capable of life-long learning.

In our earlier projects we have used such analysis both on individual system level and in cooperation with other similar systems in vehicle domains such as trucks, buses and construction equipment. In this project we plan to generalise our solutions to a new domains, namely manufacturing, district heating, and health monitoring.

The benefits from such setup come from the fact that engineers who designed this equipment are unable to predict all possible situations or conditions that they operate in. A self-monitoring system has a model, partially built by itself from the data available, of its own operation. It can evaluate its own operation and notify someone else when operation is not “normal”. An example is diagnostics, where a self-explaining system can at least partially diagnose itself and assist the human operator, as well as explain to the human operator why it came to its conclusions. This can include detecting faults that have never been thought of during design.

DISTRICT HEATING: It is estimated that 75% of buildings and district heating substations function improperly, either due to faults or incorrect configuration. These inefficiencies cause a loss of approximately 1 billion SEK per year, both to the customers but also to the network as a whole. Recently deployed smart metering systems provide large amounts of data concerning district heating operations. This data can be very valuable, but it is not possible to analyse it by hand on a large scale.

We will start by collecting data and developing the analytics methods during the first year of the project. Those will then be evaluated, and improved, based on the feedback obtained during the subsequent years. We will consider several systems in different parts of south-west Sweden in order to examine the generality of our approach, initially starting with Öresundskraft and Halmstads Energi och Miljö AB.

MANUFACTURING: In a similar manner, the intelligent agents can operate on the network level, based on HMS connectivity solutions, and hunt for "interesting" relationships or representations of the subsystem operation. They can learn models for normal operation of the connected equipment, as well as the typical interactions between them. In the project we will collect data from at least one customer using HMS networking equipment.

HEALTH CARE: A feasibility study conducted by Hälsoteknikcentrum Halland has resulted in a number of interviews in the business and public sector. A quote from the medical coordinator at the Department of Neurology and breathing at Danderyd Hospital: "A urine-sampling toilet could affect our business in that it can reduce the risk of infection among staff, offer significant time-savings, be positive for patient privacy and give us the opportunity to take more samples if necessary. Depending on what such a toilet would cost it can also affect the economy." Urine sampling is done for many different reasons by many organizations in society, such as to detect urinary tract infection, kidney problems, STDs (chlamydia) or multidrug-resistant bacteria. Within an elderly care home, it would be possible to monitor e.g. vitamin and mineral levels (potassium, nitrates, glucose, etc.) and whether the elderly take their medications as expected as well as potentially other detailed health status data of the residents in the care home.

IMPACT: The large variety of equipment types on a factory floor or customers in the district heating network or individual health and lifestyle specifics, the large number of influencing intrinsic and extrinsic factors, as well as daily and seasonal variations, make it very difficult to specify optimal operations a priori.

This project will contribute to developing advanced machine learning methods that can adapt and learn new operation patterns over time. We will develop data mining techniques that can automatically extract information about load patterns, identify malfunctioning substations and secondary systems, prioritize faults, and evaluate the effect of addressing said faults in improving operations.

Planerad projekttid -- Planned project duration

2016-02-01 – 2020-01-31

Preliminär budget (uppskatta insatsbidrag från KK-stiftelsen, ange även annan finansiering, t.ex. interna medel eller annan extern finansiär)
Preliminary budget (estimate the funding from the Knowledge Foundation; also specify other financing, e.g. internal resources or from other external funding source)

KK-stiftelsen (Knowledge Foundation)	9 500 000 kr
Högskolan i Halmstad (Halmstad University)	1 000 000 kr
Öresundskraft	1 500 000 kr
HEM	1 500 000 kr
HMS	1 500 000 kr
TBD (one of HMS customer companies)	1 500 000 kr
Villeroy&Boch Gustavsberg	500 000 kr
ACREO Swedish ICT	500 000 kr
ObserveMedical	1 500 000 kr
Volvo Group Trucks Technology	1 500 000 kr
Summa / Total:	20 500 000 kr

Ny insatsform eller avvikelse från KKS ordinarie insatsform (beskriv föreslagen ny insatsform alternativt vilka avvikelser från ordinarie insatsform som planeras)
New project form, or deviation from ordinary KKS instruments (describe the suggested new instrument / project form, or which deviations from the form that are planned)

No deviations are planned.

Beskriv insatsens strategiska betydelse för att utveckla ett eller flera av KK-miljöns forskningsteman eller på annat sätt bidra till att uppfylla de långsiktiga målen
Describe the strategic importance of the project for developing one or more of the research themes of *Research for Innovation*, or in other ways contributing to reaching the long term goals of *Research for Innovation*.

The questions raised in this project are central in our aware intelligent systems research theme. This synergy project is an extension of our research in the field of machine learning and it establishes new collaborations with local companies, allowing us to test the broad applicability of our research. It also promises to promote collaboration between IS-lab and CC-lab, in the context of network communication. It also further expands our research in the interesting and growing area of "smart grids".

The project will result in new techniques that can lead to innovative digital services and create value for the companies and their customers. Such methods also target an improvement in energy quality and efficiency, contributing to a more sustainable environment.