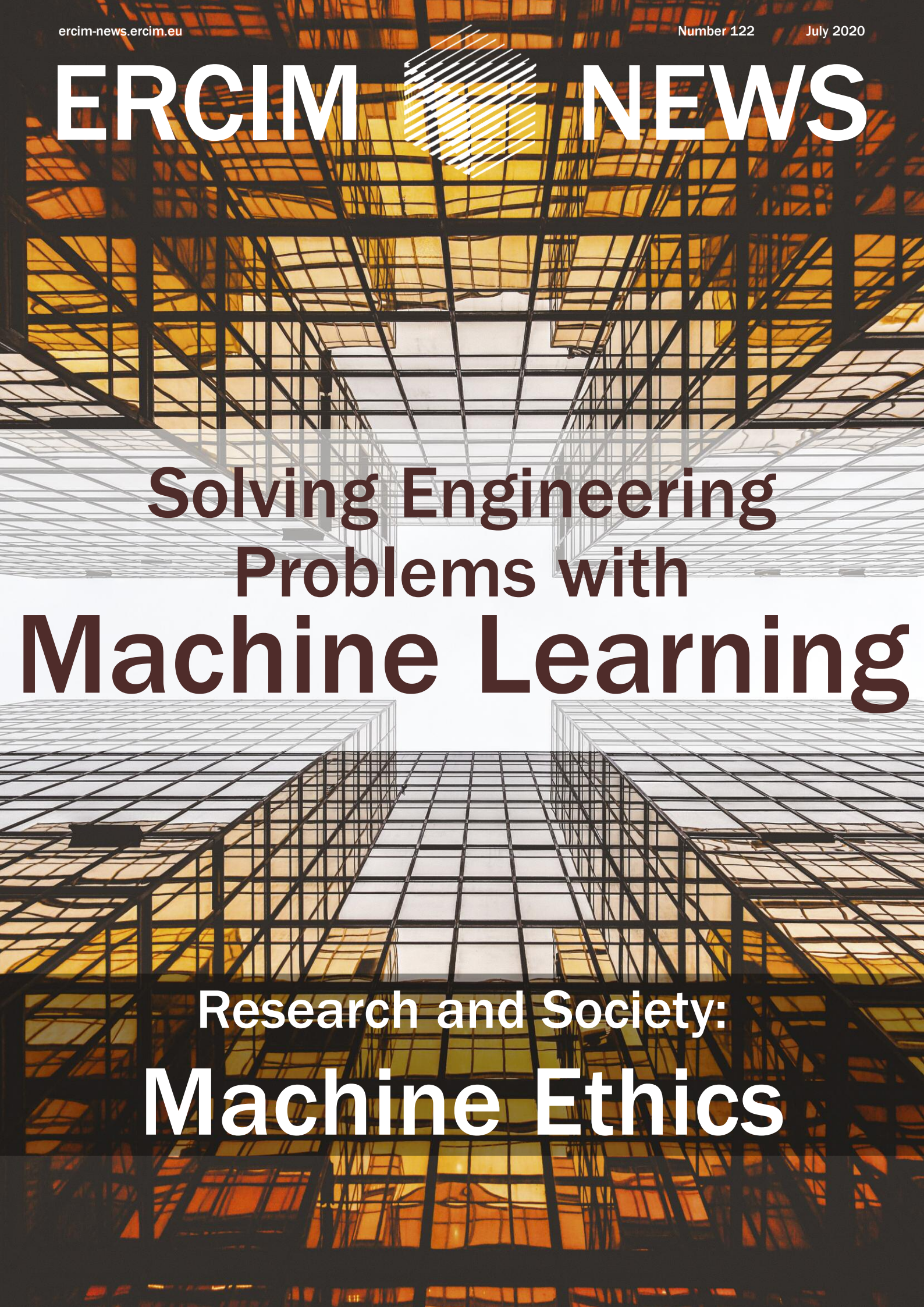
ercim-news.ercim.eu Number 122 July 2020

ERCIM NEWS



Solving Engineering Problems with Machine Learning

Research and Society: Machine Ethics

Editorial Information

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
+33 4 9238 5010, contact@ercim.eu
Director: Philipp Hoschka, ISSN 0926-4981

Contributions
Contributions should be submitted to the local editor of your country

Advertising
For current advertising rates and conditions, see https://ercim-news.ercim.eu/ or contact peter.kunz@ercim.eu

ERCIM News online edition: *ercim-news.ercim.eu/*

Next issue:
October 2020: special theme: Blue Growth

Subscription
Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu

RESEARCH AND SOCIETY

This Section with the topic “Machine Ethics” has been coordinated by Erwin Schoitsch (AIT Austrian Institute of Technology)

SPECIAL THEME

The Special Theme “Solving Engineering Problems with Machine Learning” has been coordinated by Noémi Friedman (Institute for Computer Science and Control, (SZTAKI), Hungary) and Abdel Labbi (IBM Research - Europe)

RESEARCH AND INNOVATION

ANNOUNCEMENTS

Machine Ethics

by Erwin Schoitsch (AIT Austrian Institute of Technology)

The impending highly automated and autonomous systems enabled by artificial intelligence (AI) bring with them new challenges and risks. Placing too much trust in, or misusing, machines that make decisions is risky, and the legalities are complex in terms of liability and responsibility. Autonomous systems can be grouped into three broad categories: technical systems that make decisions in "no win" hazardous situations (vehicles in traffic, collaborating robots); decision support systems in governance applications (administration, government, court, staff acquisition, etc.), which may lead to unfair decisions for humans and society; and systems that are open to deliberate misuse by providing information that can't be proven to be true or fake, potentially influencing elections, public opinion or legal processes to an extent unknown before. These risks cannot be easily countered by conventional methods. We give an overview of the potential and risks afforded by this technology.

Of course, there have long been risks associated with technology, with the potential for the dissemination of misinformation, failing algorithms and deliberate deception, but until recently the methodology at least allowed analysis and assessment of the predictable and deterministic algorithms behind the technology. We are now facing a completely different challenge – the age of highly automated and autonomous systems, artificial intelligence (AI) and decision making, whereby human decisions are made by machines through methods such as deep (machine) learning, which are neither "explainable", nor be based on fair, unbiased training sets.

Public acceptance of highly automated and autonomous systems relies on trust in these systems. This is not just a technical issue, but also an ethical one, with technology having "big brother" potential and other possible problems as foreseen in science fiction, e.g., Isaac Asimov's "Three Laws of Robotics". Asimov's laws seem reasonable and complete, but although they were complemented by an overarching "Zeroth law" ("A robot may not, through inaction, allow humanity to come to harm"), it has been demonstrated (even by Asimov himself) that realistic situations may result in unresolvable conflicts for a robot just because of adhering to this law.

AI technology is being implemented in automated driving, collaborative robots in the workspace, assistive robotic systems, highly automated production, and in management and decision systems in the medical and public service areas, the military, and many other fields. The EC, the European Parliament, the UN, many informatics and computer associations, and standardisation groups, the German Ethics Commission for Automated Driving, NGOs, and others, have created guidelines or even certificates for trustworthiness of highly automated systems, AI-systems, cognitive decision systems, automated vehicles, robotic systems, ethically aligned design, and the like (see [1], [2], [3]). A new science of "robot psychology" has evolved, that studies the interrelationship of human-robot collaboration and human wellbeing in a world of intelligent machines, and addresses how to keep human rights and individual decision making alive.

It seems that the question "Is it possible to create practical laws of robotics which can guarantee a safe, conflict free and peaceful co-existence between robots and humans?" cannot be given a definitive answer that is valid in all foreseeable situations. Even in Asimov's stories, robots had to decide which type of risk of harm is acceptable (e.g. autonomous robotic surgeon). Other authors have assumed that a mental breakdown is the logical consequence of detecting that an activity which seemed to follow Law 1 had a disastrous outcome, e.g. in "The Robots of Dawn" the story's plot revolves around a robot that was apparently destroyed by such a mental breakdown (like a "short circuit" in his computer brain).

Figure 1: Examples for Ethics Guidelines from various organizations.

These robotic laws were written in 1942, when robots were androids and just relatively simple "slaves" for humans, not the highly complex robots that are conceivable today. And what about a robot developed for an army? And who is defined as, or how do we define a "human being" (from history we know that sometimes a certain group of people is not considered as equally human and killed, e.g. genocide)? For this, we have to look at the humans behind the AI and robots. And this only partially covers the aspects of "machine decision making" and "machine ethics", referred to in the abstract.

One initiative attempting to cover the principles for system designers and developers is the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (AI/AS) (April 2016, with a document 2019 [3]) (see Figure 1). It not only identifies and recommends ideas for standards projects focused on prioritising

ethical considerations in AI/AS (i.e., machine/computer decision making), but also proposes a certificate for "ethically aligned design". The basic concept states:

"Ultimately, our goal should be eudaimonia, a practice elucidated by Aristotle that defines human well-being, both at the individual and collective level, as the highest virtue for a society…. honouring holistic definitions of societal prosperity is essential versus pursuing one-dimensional goals of increased productivity or gross domestic product (GDP). Autonomous and intelligent systems should prioritize and have as their goal the explicit honouring of our inalienable fundamental rights and dignity as well as the increase of human flourishing and environmental sustainability. The goal of The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems ("The IEEE Global Initiative") is that Ethically Aligned Design will provide pragmatic and directional insights and recommendations, serving as a key reference for the work of technologists, educators and policymakers in the coming years."

Many standardization groups, the EC HLEG group document [1], and the German Ethics Commission on Automated and Connected Driving [2], provide a set of recommendations for decision making, placing human rights, independence and wellbeing in the centre, independent of economic or demographic attributes, such as age and race. But within the Trustworthiness groups on ethics and governance in ISO/IEC JTC1 SC41 (IoT) and SC42 (AI), the international discussion revealed that even the definition of (individual) human rights differs among cultures and different countries' legal systems.

Perhaps it makes sense, then, to focus on "easier" issues in our cultural environment. The article "Machine Learning Based Audio Synthesis: Blessing and Curse" reveals an important risk: The benefits of helping people with special needs to express themselves verbally are likely to be marginal compared with the risk of eroding trust in audio media, with the potential for "deep fakes" to become this technology's major application. "Covering Ethics in CPS Design" reports about the "Civis 4.0 Patria" project, on citizen participation in disaster management, which relies on trusted information and a trustworthy framework (cybersecurity, privacy, safety as some of the main properties), with an approach leading to an ethically aligned software design and security architecture.

"Trustability in Algorithmic Systems Based on AI in the Public and Private Sector" addresses the governance challenge of machine-driven decision making in public administration, criminal justice, education and health. In addition to technical safety solutions, trustability and interpretability are key issues. Legislation, fairness and ethical principles (which, in certain contexts, contradict themselves) are main concerns. These systems should be evaluated by the public and checked for acceptance based on their decisions and the public's perception of them. The "human comprehensible model" as addressed in the paper is a key precondition.

The article "Why your Robot Co-Worker Needs a Psychologist" addresses the new interdisciplinary research challenge of robot psychology. It outlines the research and development needs for this area, as addressed in the Austrian research lab "CoBot Studio", where experts from different disciplines work together towards the common goal of human-centred trustworthy collaborative robots.

The article "You Can Make Computers See; Why not People?" discusses the achievements that have been made in the areas of computer vision and image recognition, e.g. for highly automated vehicles and storage control. The article queries why these technologies have not been more widely used to help vision-impaired people ("assisted vision"), and raises the question of whether our efforts might be better spent trying to directly help humans.

The issues raised in these articles represent only a few of the big dilemmas that we need to address. Hopefully the articles in this section will motivate the reader to ponder the ethical questions raised by technologies, their design, implementation and application.

Part of the work described received funding from the EC (Horizon 2020/ECSEL Joint Undertaking) and the partners National Funding Authorities (in Austria the Austrian Research Promotion Agency (FFG) and the Federal Ministry for Climate Action, Environment, Mobility, Innovation and Technology (BMK)) through the projects AutoDrive (737469), Productive4.0 (737459) and SECREDAS (783119).

Link:

[L1] "When computers decide" - Informatics Europe and ACM Europe: https://kwz.me/hEE

References:

[1] European Commission, Independent High-Level Expert Group, "Ethics Guidelines for Trustworthy AI" (Final report April 2019, HLEG AI), Brussels. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[2] Federal Ministry of Transport and Digital Infrastructure, Ethics Commission on Automated and Connected Driving – Report June 2017, Germany; Summary available in English on https://kwz.me/h0O

[3] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems", First Edition. IEEE, 2019. https://kwz.me/h0k

Please contact:

Erwin Schoitsch
AIT Austrian Institute of Technology and Secure Business Austria (SBA)
Erwin.schoitsch@ait.ac.at

Machine Learning Based Audio Synthesis: Blessing and Curse?

by Nicolas Müller (Fraunhofer AISEC)

Machine learning based audio synthesis has made significant progress in recent years. Current techniques make it possible to clone any voice with deceptive authenticity, based on just a few seconds of reference recording. However, is this new technique more of a curse than a blessing? Its use in medicine (restoring the voice of the mute), in grief counselling or in the film and video game industry contrasts with its enormous potential for misuse (deep fakes). How can our society deal with a technology that has the potential to erode trust in (audio) media?

Using artificial intelligence, voices can even be reproduced through short audio sequences. (Source: Fraunhofer AISEC).

Owing to groundbreaking success in various disciplines, artificial intelligence (AI) is on everyone's lips, both figuratively and literally: using deep neural networks, researchers have recently developed a system that can reproduce any voice in a deceptively realistic way. Using only a few seconds of reference audio material, the AI recognises the characteristics of the person speaking, can reproduce the voice accordingly and thus place any sentence in the person's lips.

This is definitely a technical masterpiece with a variety of use cases. For example, it would allow people who have lost their voice due to accident or illness to communicate with a replica of their natural voice via a human-computer interface. Similar scenarios are also imaginable in grief counselling. Nevertheless, there are also great opportunities for artists and cultural workers and for the film and video game industry. Finally, this technology is a cornerstone for a new, more human form of artificial intelligence that will accompany us in our daily lives even more intensely than current systems.

This technology is already making impacts: Scientists have developed a system that can exchange single words in any sentence. Like its well-known image editing namesake, "Photoshop for Voice" allows free editing of audio [L1]. During a demonstration [L2], for example, the sentence "I kissed my wife" was changed to "I kissed Jordan".

It is also possible to create entire sentences, resulting in a perfect illusion, providing the lip movements are synchronised accordingly. For example, in a well-known Deepfake video clip [L3], Boris Johnson supports his political opponent Jeremy Corbyn and even recommends him as Prime Minister of England. These examples are created without malicious intent, but at the same time, they illustrate how easily this technology can be misused.

As a result, it is obvious that in the future we will not be able to trust audio and video material unconditionally. Fake news, i.e., deliberately created false information, is continually improving, making it even easier to deceive us. In the US election of Donald Trump in 2018, voters were manipulated massively by fake news [L4]. This influence is likely to increase and machine-learning based audio synthesis may be used to defame political opponents.

Facing this challenge will be a central goal for the coming years. How are we supposed to deal with it? Banning the technology is simply not feasible. Since basic source code [L5] and technical specifications [L6] are already public, such AI systems will be widely available for anyone to use. It will not be a technology dominated by a few.

Therefore, what remains? One way to strengthen the trustworthiness of digital media could be a second AI: an AI that is an expert in distinguishing between generated and real audio material. This could include certificates of authenticity or warnings for counterfeit material. The Fraunhofer Institute for Applied and Integrated Security AISEC is currently actively researching AI systems that, as experts, can differentiate between genuine and fake media content. Such systems learn this distinction by using a training data set containing a large number of both "real" and "fake" audio examples. In this way, the expert AI learns to detect subtle irregularities in the fakes, which are not noticeable to humans but are nonetheless present. Thus, deep fakes can be detected.

Yet, we will have to learn to rethink: Just as we should not trust every article published on the Internet, we must start to mistrust every audio or video clip. Because it is not a question of "if", but of "when" fake material will affect us on a large scale. When the time comes, we should be prepared both technically and socially.

Links:

[L1] https://kwz.me/h0s
[L2] https://www.youtube.com/watch?v=I3l4XLZ59iw
[L3] https://www.youtube.com/watch?v=30NvDC1zcL8
[L4] https://www.nature.com/articles/s41467-018-07761-2
[L5] https://github.com/CorentinJ/Real-Time-Voice-Cloning
[L6] https://arxiv.org/abs/1806.04558

Please contact:

Nicolas Müller, Fraunhofer AISEC, Germany
+49 89 3229986 197, nicolas.mueller@aisec.fraunhofer.de

Covering Ethics in Cyber-Physical Systems Design

by Christoph Klikovits (Forschung Burgenland), Elke Szalai and Markus Tauber (FH Burgenland)

Digitisation is leading to the increased use of cyber-physical systems (CPS). A citizen participation and disaster management platform uses IoT components like sensors, which collect information about critical events in disaster scenarios. In this situation it is critical that all stakeholders can be assured of trustworthy information. We are researching an approach that takes ethical considerations into account during the development process, resulting in a secure, trustworthy framework.

Increasing, due to availability of technology, also smaller communities make use of the automation of processes and involvement of citizens, creating "smart municipalities", those rely. Internet of Things (IoT) components are used to improve processes of the local administration and communication. Most current platforms and applications have a restricted point of view, with a focus on features and sustainability [1]. To achieve trustworthiness and secure communication—which are essential but often neglected in systems rely on citizen participation—ethical aspects of the system must be considered. Ideally, all stakeholders should have input and these considerations factored into the system early in the design process.

In the EFRE project (FE07) "Civis 4.0 Patria" [L1] a citizen participation and disaster management platform will be developed. The project aims to make it easy for citizens and local authorities to communicate with each other and to share information [1]; for example for citizens to report incidents, such as open manholes or potholes, to the local authorities. Furthermore, local IoT sensors and external weather services provide disaster warnings and can improve the lead time to achieve an accurate operational picture. This information can be processed by local authorities, citizens, or emergency services. IoT sensors will collect information about air quality, water levels, rainfall, temperature, storm, heat, fill levels and number of visitors, which can give an overview of the situation to emergency personnel.

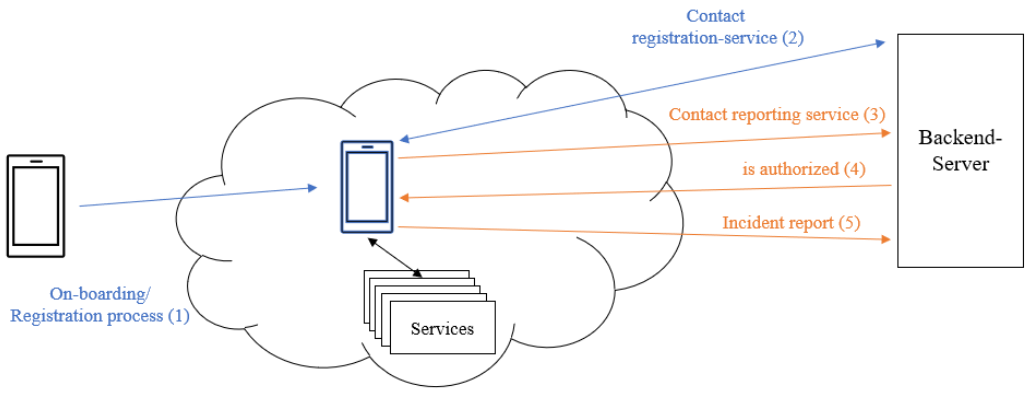
Disaster management and citizen participation are sensitive areas in which privacy and trustworthiness are important. The stakeholders determine what action to take based on the information they receive, and this works in both directions: from the citizens to the public administration and vice versa. Hence, trustworthiness of information and the safety of individuals are paramount.

We are researching the integration of ethical principles during the design phase of a citizen participation and disaster management platform. This approach can lead to technology acceptance by users.

To ensure that the needs of different groups of stakeholders are considered, we incorporate social science methods, such as the use of "personas", during the development process. This process gives stakeholders the opportunity to express worries, doubts, objections and suggestions. The "persona" method makes it possible to describe different types of user, to understand their needs, and to consider them when making design-related decisions. In a first step, the research team will develop "personas", which result from interviews and questionnaire surveys. These personas represent prototypes of the project and encompass stakeholders' ethical concerns about the system's trustworthiness and security. Features addressing these concerns can be proactively build into the software design and security architecture. This approach increase user acceptance of the technology and can help researchers recognise and perhaps change the attitudes of users [3].

Once the platform is developed, citizens, local authorities, and others can register and use it for purposes such as reporting incidents with their smartphones (for example). To this end, we need to identify a secure IoT framework that can integrate smartphones, sensors, and external services into the platform while achieving trustworthy, secure communication and safe storage of data.

The Arrowhead framework [L2] and its onboarding procedure is one method that we are considering for this purpose. The Arrowhead chain of trust [3] enables a trustworthy environment through its process-oriented usage of certificates and secure onboarding of smartphones. The registration and incident reporting processes are researched in the opensource



framework Arrowhead, which provides many security functions by design. First, smartphones can be onboarded in the Arrowhead framework, as shown in Figure 1. They form part of the Arrowhead local cloud, authorised by certificates, and can share information in a trustworthy way, e.g., reporting incidents in the context of "Civis 4.0 Patria".

In contrast to hardware components (e.g., smartphone or sensor) being onboarded, it is not only the device but also the user that governs the interactions and hence new approaches for onboarding in cyber-physical systems may be needed. The researched techniques (Arrowhead onboarding and personas) represent a possible approach with the objective to facilitate security, safe storage of data and trustworthiness, in convergence with ethical considerations within a citizen participation and disaster management platform. The incorporation of personas allows the needs of individual stakeholders to be factored in early in the development process. This approach leads to ethically aligned software design and security architecture.

Links:

[L1] https://www.forschung-burgenland.at/it/civis-40-patria/
[L2] https://www.arrowhead.eu/

References:

[1] J. Wolfgeher, M. Zsilak, M. Tauber: "Smart Municipality", ERCIM News, issue 119, 2019, https://ercim-news.ercim.eu/en119/special/smart-municipality
[2] A. Bicaku, et al.: "Interacting with the arrowhead local cloud: On-boarding procedure", in 2018 IEEE Industrial Cyber-Physical Systems (ICPS), pp. 743-748. IEEE, 2018.
[3] J. Pruitt, J. Grudin: "Personas: practice and theory", in Proc. of the 2003 conference on Designing for user experiences (DUX '03), ACM, 1–15, 2003. https://doi.org/10.1145/997078.997089

Please contact:

Christoph Klikovits
Forschung Burgenland, Austria
christoph.klikovits@forschung-burgenland.at

Trustability in Algorithmic Systems Based on Artificial Intelligence in the Public and Private Sectors

by Sónia Teixeira, João Gama, Pedro Amorim and Gonçalo Figueira (University of Porto and INESC TEC, Portugal)

Algorithmic systems based on artificial intelligence (AI) increasingly play a role in decision-making processes, both in government and industry. These systems are used in areas such as retail, finances, and manufacturing. In the latter domain, the main priority is that the solutions are interpretable, as this characteristic correlates to the adoption rate of users (e.g., schedulers). However, more recently, these systems have been applied in areas of public interest, such as education, health, public administration, and criminal justice. The adoption of these systems in this domain, in particular the data-driven decision models, has raised questions about the risks associated with this technology, from which ethical problems may emerge. We analyse two important characteristics, interpretability and trustability, of AI-based systems in the industrial and public domains, respectively.

Data-driven models, as well as other algorithmic systems based on artificial intelligence (AI), must be in accordance with legislation and regulations, upholding the law, values and ethical principles. Unfairness is one of the greatest ethical concerns from the emergence of technological risks. In addition, it is necessary to guarantee the right to explanation, instituted with the GDPR (General Data Protection Regulation). In areas of public interest that use algorithmic systems based on AI, namely, data-driven decision models, we intend that these be fair, effective and transparent when it comes to ensuring the right to explanation.

In this thesis project, three dimensions of risk are being considered from a technological perspective (bias, explainability and accuracy), and two from an ethical perspective (fairness and transparency). The AI process consists of several phases, and it is important to determine whether all phases have the same type of risks. If they do not have the same type of risks, it is important to understand where each type of technological risk under analysis is located in the process. This identification was our first stage of work [1], which currently allows us to easily identify which phase requires our attention when we intend to solve a specific type of risk (whether from the perspective of AI, or from the perspective of political decision-makers). Figure 1 shows our framework for a more trustable system, in the public interest domain.

Our approach considers the interconnection between the three dimensions of risk from a technological perspective (bias, explainability and accuracy) together with the ethical

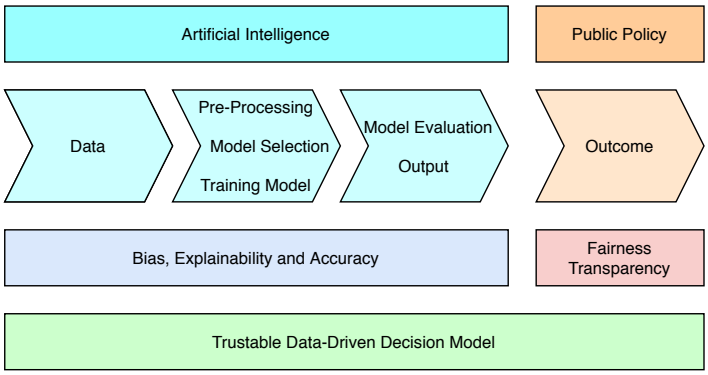


Figure 1: Framework of the risks involved for a trustable system in areas of public interest.

risks, all underpinned by our values and law, to obtain more ethical and trustable algorithmic systems. Data-driven decision models can have a bias at several stages of the process, which can contribute to injustices. For this reason, both explainability and the right to explainability are important, in order understand and explain the decisions made by the system. However, sometimes the models are so complex that it is not possible to know whether wrong associations were learned and whether the decision support provided by the system is as good as the accuracy indicates. At the moment we are selecting the most suitable case-study from these hypotheses.

Algorithm/model explainability remains an important unsolved problem in AI. Developing algorithms that are understandable by factory workers, for example, could substantially boost the deployment of such systems. Hence, Explainable AI (XAI) is an emergent research field that may drive the next generation of AI. Most of the existing approaches to tackle XAI start by applying a black-box model (BBM) to the problem at hand. BBMs, such as neural networks and tree ensembles, are well-established machine learning approaches that usually perform well. In a second step an explanator is built to help interpret the results of the BBMs. However, this two-step approach of first learning the BBM and only then building the (global) explanator, has important drawbacks: i) much of the fine-tuning of the first step is lost in the second step, as the explanator does not absorb it all; ii) the explanator will have the biases of the BBM.

Our proposal is to learn directly, in a single, joint phase, a human comprehensible model. This approach will not be possible in every machine learning application, such as those using image or text data (which is predominantly done with deep learning and increasingly already trusted and explainable), but there is wide class of problems that use tabular data, including classification/regression and that leverage this concept. Fundamentally, we propose the use of symbolic learning methods that work with analytical expressions that humans can understand and improve on (c.f. Figure 2).

The project regarding trustability in data-driven decision models in the areas of public interest is part of a PhD thesis in Engineering and Public Policy, started in 2019, at the Faculty of Engineering of the University of Porto, with the host laboratory the Laboratory of Artificial Intelligence and Decision Support from INESC TEC, and is supported by Norte Portugal Regional Operational Program (NORTE 2020). The project focussing on interpretability TRUST-AI project, supported by the EU Commission, is starting this year.

Reference:

[1] S. Teixeira, J. Rodrigues, J. Gama: "The Risks of Data-Driven Models as Challenges for Society", in IEMS '20 — 11th Industrial Engineering and Management Symposium: The Impact of DEGI Research on Society, Porto, Portugal, 51-53, 2020.

Please contact:

João Gama, INESC TEC, Portugal
jgama@fep.up.pt

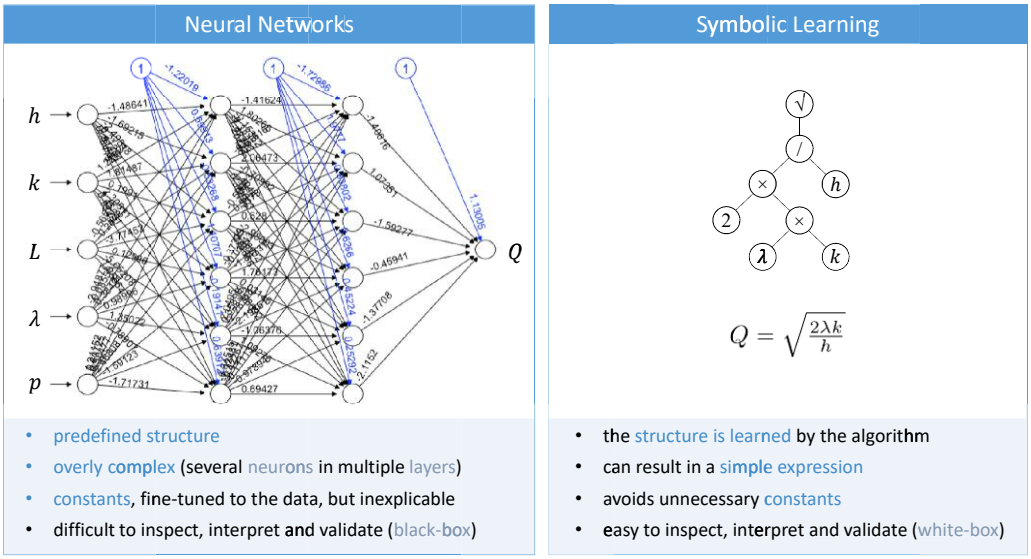


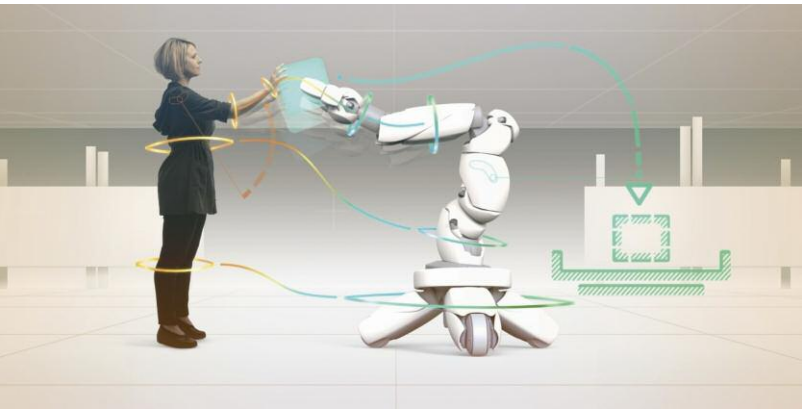
Figure 2: Comparison of a BBM (neural network) and a white-box model (obtained via symbolic learning).

Why your Robot Co-worker Needs a Psychologist: Interdisciplinary Research for Trustworthy Machines

by Martina Mara (Johannes Kepler University Linz)

As close collaborations between humans and robots increase, the latter must be programmed to be reliable, predictable and thus trustworthy for people. To make this a reality, interdisciplinary research and development is needed. The Austrian project CoBot Studio is a research initiative in which experts from different fields work together towards the common goal of human-centred collaborative robots.

Trust is a fundamental building block of relationships. This is true not only for personal life, but also for relationships at work. People who cannot trust their co-workers are likely to feel insecure, be less efficient and experience less job satisfaction. A look into the emerging field of collaborative robotics reveals that, at least in some work environments, these co-workers will increasingly be robots.



In the interdisciplinary research project Cobot Studio, a virtual simulation environment for trustworthy human-robot collaboration is being developed.

Unlike conventional industrial robots, collaborative robots—or CoBots for short—are light, safe and intelligent enough to operate in close physical proximity to people. CoBots will be increasingly used in production halls, warehouses and healthcare, working side by side with employees to conduct tasks such as assemble car seats, inspect packaging and prepare medication. Industrial robots have typically been put behind barriers or in cages for safety reasons, minimising their interactions with humans. The concept of the CoBot changes all this: the formerly isolated industrial robot becomes a social machine, sharing its space with humans and thus requiring an understanding of the states and goals of people. CoBots must be programmed in such a way that they can be trusted by their co-workers.

But what does it actually mean to trust someone (or something)? Psychologists have been concerned with this question for a long time. According to the American Psychological Association, the largest scientific organisation of psychologists worldwide, trust refers to the confidence that one party has in the reliability of another party. More specifically, it is described as the degree to which one party feels that they can depend on another party to do what they say they will do. Therefore, a key factor for the attribution of trustworthiness to a person (or a robot) is this person's (or this robot's) predictability. Being able to anticipate whether an interaction partner will act in accordance with one's own expectations is considered essential not only in the scientific literature on cognitive trust in interpersonal relations [1], but also in the Trust in Automation literature [2].

Applied to the area of human-robot collaboration, this means: Just as the states and intentions of the human partner must be identifiable for the robot, so too should the states and planned actions of the robot be easily understandable and predictable for the human co-worker [3]. This becomes especially important in situations that psychologists describe as trust-relevant, i.e., situations that involve vulnerability and risk, whether social, financial, personal, or organisational in nature. Trust is also thought of as something procedural. As the number of (positive) experiences with an interaction partner increases—let's say because an employee has been working successfully with a particular robot on similar tasks for a long time—the perceived predictability and thus trustworthiness should naturally increase along with them. However, one vision associated with CoBots is that even non-experts or employees who have not undergone any special training should be able to use them. Therefore, trustable CoBots must signalise their intentions in a manner that is also understandable for people with limited or no previous experience. For instance, when a CoBot is about to actively intervene in a work process, its actions, such as which direction it will move and which object it will grip, should be intuitively apparent to nearby humans.

But which signals and interfaces are the best indicators of where a robot is about to move or what it is going to do next? This question still needs addressing. The development of easily understandable intention signals in CoBots and the empirical evaluation of their assumed association with trust and acceptance from the perspective of the human co-worker requires research that is characterised by an interplay of many different disciplines and their complementary views on the topic. CoBot Studio is an exemplary collaborative research endeavor that gathers experts from various fields, including robotics, artificial intelligence, psychology, human-computer interaction, media arts, virtual reality and game design around a shared vision of "Mutual Understanding in Human-Robot Teams". Funded by the Austrian Research Promotion Agency FFG and running from 2019 to 2022, the project focuses on the development of an immersive extended reality environment in which collaborative tasks with mobile industrial robots are simulated and the effects of different light-based and motion-based intention signals conveyed by these virtual robots can be studied under controlled conditions.

Using state-of-the-art VR headsets, study participants in the CoBot Studio play interactive mini games in which tasks such as organising small objects together with a CoBot or guessing the target location of a moving robot in space have

to be completed. During the games, the robot's intention signals (e.g., LED signals or nonverbal communication cues) are varied and their respective impacts on comprehensibility, trust, perceived safety, and collaborative task success are evaluated. After several iterative runs of such virtual CoBot games, findings about the effectiveness of different intention signals will be evaluated in a game with physically embodied CoBots. Based on this method, the interdisciplinary CoBot Studio team intends to create practice-oriented guidelines for the design of predictable and thus trustworthy collaborative robots.

Links:
http://www.cobotstudio.at/
https://www.jku.at/lit-robopsychology-lab/

References:

[1] J. K. Rempel, J. G. Holmes, M. P. Zanna: "Trust in close relationships", Journal of Personality and Social Psychology, 49, 95-112, 1985.
[2] J. D. Lee, K. A. See: "Trust in automation: Designing for appropriate reliance". Human Factors, 46, 50-80, 2004.
[3] A. Sciutti, et al.: "Humanizing human-robot interaction: On the importance of mutual understanding", IEEE Technology and Society Mag., 37, 22-29, 2018.

Please contact:

Martina Mara, LIT Robopsychology Lab, Johannes Kepler University Linz, Austria, martina.mara@jku.at

You Can Make Computers See; Why not People?

by Anna Leida Mölder (NTNU)

What is it that motivates researchers to advance knowledge in a particular field? What is it that makes an engineer develop a tool for one specific purpose, but not another? And how do these choices of today affect the future?

In the era of artificial intelligence (AI), one of the most researched topics is how to teach computers to see, and to implement this computer vision in a range of applications, such as drones, microscopes and cars. At the same time we have at least 2.2 billion people in the world suffering from some type of vision impairment [L1].

Many technological advances have been made to help people with visual impairment. Voice navigation is available for most map functions and a number of mobile phone applications exist, such as iMove around [L2] and RightHear [L3]. But just like vision technology in cars, knowing where to go and how to get there is not enough. Users must also respond to dynamically appearing changes in their surroundings. In self-driving cars, intelligent systems based on a combination of sensors, such as LIDAR, cameras and time of flight measurements have been implemented to prevent the vehicle from running into people, animals or other cars. Depth-sensing cameras are being further developed and used recreationally by some of the world's largest corporations. But for people with vision impairments, the go-to assist for depth sensing is still a guide dog or a white cane.

While there is nothing wrong with a guide dog—which is probably cheaper, more accurate and a better companion than any self-driving car—how can we explain the strong bias of technological development? A search on Google Scholar for "vision impairment LIDAR" turns up 7,220 hits. A search for "car LIDAR" turns up 126,000 and a search for "self-driving car" 2,760,000. This represents a more than 10-fold increase in the quantity of research being undertaken for the same technology for two very different purposes.

Image recognition and identification are also increasingly accessible for mobile devices, with the construction of less complex AI models, such as MobileNet [1]. It is easy to find applications to classify anything from flowers to car number plates and cat and dog breeds, using smart phone captured images. The same technology could be used to help people with vision impairments identify specific items in their surroundings; things that the seeing community takes for granted. Imagine being able to shop for food and actually being able to find the Fair Trade logo on the package, being able to read the "best before" date or to figure out if the product contains the lactose you are allergic to, without having to ask for assistance every time.

QR codes were once developed especially with the vision impaired in mind. They are easy to access via a handheld device, which can in turn be configured to deliver the output by any means that the user chooses, audio or on screen. They can even be used to help customers make phone calls or send messages to a support function or store owner—a great help for any person with vision impairments entering to do their shopping. However, very few retailers place QR-codes on their products for any purpose other than promotion or discount management. Blind people often use Braille tags to identify items of clothing in their wardrobe. Similarly, retailers could use existing technology to help people with vision impairments and make the store accessible to all customers.

In this era of computer vision, facilitated by improved mobile device resources, assisted vision for many humans is lagging. As AI encroaches into our daily lives, it is important that we consider who we include in this new technological era, and how. Many applications will never be developed for the simple reason that researchers, engineers and developers are simply not spending their time working on it right now—today. Are we all really doing the best we can be doing, every day?

Links:
[L1] https://kwz.me/h0a (8 Oct 2019)
[L2] http://www.retinaitalia.org/imove-around-per-android/
[L3] https://www.right-hear.com/

Reference:

[1] A. G. Howard et al.: "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017. https://arxiv.org/abs/1704.04861

Please contact:

Anna Leida Mölder, NTNU, Norway
anna.l.molder@ntnu.no

Introduction to the Special Theme

Solving Engineering Problems with Machine Learning

by Noémi Friedman (Institute for Computer Science and Control, (SZTAKI)), and Abdel Labbi (IBM Research - Europe)

Machine learning (ML) brings many new and innovative approaches to engineering, improving efficiency, flexibility and quality of systems. This special theme of ERCIM News focuses on ML applications in industrial engineering (see keynote by Christopher Ganz on page 14), with a focus on civil, environmental, mechanical, chemical, process, agricultural, and transportation engineering.

Data-driven, surrogate and hybrid modelling

ML models can learn in a progressive manner from empirical evidence, which makes them great candidates for continuously correcting modelling errors and adapting to drifts in engineering systems, that are classically modelled by the laws of physics. Surrogating and/or combining these simulations with ML algorithms can overcome limitations of knowledge and computational capacity and lead to improved predictions and engineering innovations (see the introductory paper on the advantages of ML-based surrogate modelling, Asgari et al., p. 16).

As reflected in this special theme, data-driven modelling, surrogate modelling and hybrid modelling (a combination of physics-based and learning-based models) have been successfully used in various engineering applications.

Digital twins or meta-models that replace computationally expensive physics-based simulation models can significantly reduce the computational time of modelling complex systems, such as the simulation of a methanation reactor in a power-to-gas process (Asgari et al., p. 16) or the dynamic simulation of a tall timber building (Kurent et al., p. 29).

Complex and highly non-linear systems can be extremely sensitive to changes in their governing parameters. To quantify the uncertainties of model outputs due to the possible deviations of the input parameters from their estimated value, one is often forced to run instances of the deterministic simulation model over a wide range of parameters. The same is true when parameters have to be identified, calibrated, or optimised. The computational time of ML-based surrogate models can be a small fraction of that of the original physics-based model and so they can enable an efficient way of handling such problems (Hoang et al., p. 23). Some examples presented in this special theme include the estimation of aerofoil aerodynamic performance statistics (Liu et al., p. 20), the surrogate-based calibration of tall timber building dynamics (Kurent et al., p. 29), the reduced order flow simulation by a theory and data driven hybrid model (Deng et al., p. 30), and the deep learning model for an accurate spatio-temporal rainfall estimation (Folino et al., p. 24).

ML algorithms can also contribute to coarse-grained models. Coarse-grained models are simplified models usually defined on a coarse grid or scale that can accurately simulate the phenomena that happens on a finer scale. Van Halder et al. (p. 27) describe the process of creating a coarse-grained model that simulates the sloshing motion of water, and Karavelic et al. (p. 21) describe a probabilistic scale bridging of micro- and macro-scales to model the plastic behaviour of heterogenous composite materials. Both models apply ML tools for upscaling.

ML algorithms have been receiving particular attention in the field of autonomous vehicles. Using sensor data, learning agents can address the problems of traffic congestion, energy consumption and emissions. Nevertheless, when it comes to passenger safety, a learning-based control design can never give a 100% guarantee of avoiding emergency scenarios. In such cases, a hybrid model that combines the benefits of model-based and learning-based control design can provide an efficient but still robust compromise (Németh et al., p. 18).

ML learning tools can control not only the autonomous vehicle but also the flow around it, enabling a more efficient vehicle design. Using sensor data and actuators, an automatic ML-based closed loop control can be built (Cornejo-Macedas et al., p. 32) to reduce drag or to increase lift, usually by aiming to avoid flow separation. Manipulating the flow in this way can increase performance and reduce energy consumption – important goals in aircraft design.

ML for production control and process optimisation

Process optimisation aims to reduce production time, optimise material and energy consumption, and increase product quality. Manufacturers can profit greatly from ML tools that can discover hidden dependencies between production parameters, foster production efficiency and flexibility, and manage complex optimisation tasks (Samsonov et al., p. 38).

ML tools are always based on observed data, which, unfortunately, is often difficult or expensive to collect or may raise privacy concerns. The more data available, the more ML can discover and improve. We may substitute for collecting additional data by synthetic data generation, which is called a soft

or virtual sensor. Garcia-Ceja et al. (p. 43) describe a soft-sensing system for the optimisation of a chemical process.

In cases where synthetic data cannot replace actual data, we may still improve prediction accuracy by incorporating all available background engineering knowledge. For example, an agricultural prediction model of yield of nitrogen status can be improved by combining ML tools with complex systems theory (Raubitzek et al, p. 44).

The high dimensionality of descriptive data can cause another type of problem for process optimisation (Savvopoulos et al., p. 41, Gaudin et al., p. 45). Autoencoders enable high dimensional data to be encoded in a much smaller dimensional representation, and optimisation tasks can be carried out in this reduced latent space. The use of a variational autoencoder - one that learns a probabilistic rather than a deterministic description of the latent variables - can increase the robustness of the description (Gaudin et al., Savvopoulos et al.).

Switching from a deterministic to a probabilistic approach can also ameliorate the so-called inverse problems, in which the input parameters of processes or models are to be calibrated or optimised. Inverse problems are usually ill-posed, since several values of the parameters may result in an equally good fit for the desired or measured output of the model or process. Consequently, a probabilistic description of the optimised or calibrated parameters (Hoang et al., p. 23, Smaragdakis et al., p. 35, Kurent et al., p. 29) gives a more robust and informative solution.

Monitoring and anomaly detection

Inherent changes in the environment and the system itself can create anomalies or drifts that incrementally build up and result in performance degradation. Continuous monitoring and control are therefore essential for the optimal operation of most engineering processes and systems.

The use of modern machine learning methods, such as Deep Learning and Graph Neural Networks allows complex system behaviour modelling without the need to define a large, and usually partial, set of rules and patterns of "normal" behaviour that can quickly become obsolete with time. The application of machine learning methods is made possible by the extensive instrumentation of most engineering systems and processes as well as the high frequency at which those sensors operate. This leads to innovation in monitoring (such as the new positioning system using 5G millimetre wave networks (Gante et al., p. 26) and renders classical feature-based and rule-based methods for anomaly detection obsolete because of the combinatorial amounts of data and feature combinations and rules. Deep Neural Networks are particularly well suited in such cases as they automatically learn a reduced dimensionality representation in which anomalies are more efficiently characterised (Kumar Jha et al., p. 47). By continuously retraining the system as new data arrives, machine learning models continuously adapt their internal representation of the normal and anomalous behaviour.

This special theme features several examples of combining machine learning techniques and domain-specific models for monitoring and optimising systems in domains such as autonomous transportation (Sahin et al., p. 51; Lo Duca et al., p. 52) and neonatal intensive care (Földesy et al., p. 50).

Innovative approaches to anomaly detection in complex networks are addressed by Gutiérrez-Gómez et al. (p. 49) in the context of anomalies or outliers of node attributes in graphs. Defining anomalies in a subgraph or a view is an interesting approach to multi-context anomaly detection, since graphs can represent complex dynamics that are difficult to characterise at a global scale.

As fine-grained instrumentation becomes pervasive in system and process engineering, the adoption of machine learning methods for data-driven anomaly detection, monitoring, and online control is becoming mainstream engineering of complex systems.

Conclusion

This special theme of the ERCIM News explores different fields in which ML algorithms are replacing or enhancing analysis-based methods. By using all available data to simulate complex engineering systems, we can reduce computational time or increase accuracy and efficiency. ML can help engineers create designs with increased performance and reduced consumption, identify hidden dependencies and anomalies, and optimise and control manufacturing. Nevertheless, a knowledge gap still exists between engineering, manufacturing, and big data analysis. We strongly encourage initiatives to close the gap as described by Bernijazov et al. (p. 36), and improve the efficiency of ML tools as outlined by Muccini et al. (p. 33) and Pikoulis et al. (p. 39).

Please contact:
Noémi Friedman
Institute for Computer Science and Control, (SZTAKI), Hungary
n.friedman@ilab.sztaki.hu

Abdel Labbi
IBM Research - Europe, Zurich, Switzerland
abl@zurich.ibm.com

Keynote

Machine Learning in Engineering - A View from Industry

by Christopher Ganz (ABB Future Labs)

In recent years AI has rapidly developed across many fields, finding its way into applications where it hasn't succeeded previously, and reaching into areas that were unthinkable even a few years ago.

AI and machine learning (ML) have found their place in industry. With machine learning showing its particular strength in areas that map an input data set to an output set or conclusions, its predominant applications are becoming those of classification or perception. Current success stories around ML applications often focus on condition monitoring: determining the current health status of an asset, based on a given set of measurements. The inclusion of data that led to a failure not only allows diagnosis of failures, but also prediction of failures, giving the operator a chance to reduce or stop production and resolve the issue.

Perception applications that have been widely used in consumer goods (e.g., image, video and voice recognition and natural language processing) are now being adapted to industrial settings. Whether it is video-based quality control, voice-controlled equipment, or other means of assessing plant status through sound, image, or video analysis, consumer algorithms are improving the overall performance of the plant in industrial settings.

In many cases, key performance aspects are not defined during operation, but earlier in the design process: the engineering phase. Design decisions influence the subsequent performance of a plant.

Industrial value chains

Process industries vs. discrete

Figure 1 shows a simplified view of the value chain, from raw material extraction through to a finished product in the hands of the consumer. The requirements for the engineering process to build the plants along that value chain vary. In raw material extraction and process industries, as well as utilities, continuous production of material prevails. Material is mostly transformed in a continuous process, which in some cases can run uninterrupted for days, months or even years. The production process varies little. Required adaptations, largely due to input material variations or variants produced, can mostly be handled through parameter variations in the production process.

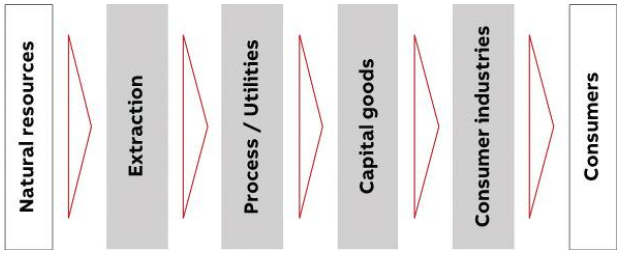


Figure 1: Value chain across industries.

The process generally remains as engineered for very long periods, even for the entire plant lifecycle.

As the production process progresses towards the customer, it moves from a continuous flow of material to the production of individual devices (discrete manufacturing) in distinct production steps. Capital goods are mostly built to order based on customer specifications, and consumer products vary along fashion trends, prompt introduction of new technologies, and other driving factors that require frequent changes in production. This may include changes in the production process that require new production technologies with corresponding new machines. In the engineering phase, the plant is given the flexibility to react to variations in the process, but very often the process is (partly) re-engineered when new products enter production.

Plant construction value chain

Engineering as we cover it here has many different facets. From the moment a company decides to produce a particular range of products to the point at which they emerge from the factory is a long journey, involving many stakeholders. Many companies do not build the factory themselves; they employ an EPC (engineering procurement construction) company to design the factory and procure the machines that go into it. The machine builder then orders components from another supplier. The equipment that makes up a plant is a hierarchical structure of systems and sub-systems, that all interact to fulfill the purpose of the plant. The engineering steps involved along the plant value chain are quite diverse. System integration engineering at plant level requires different skills and methods from the design of a product, e.g. a motor. Such a motor, once designed, is again produced in a factory, that in turn was built earlier using the very same process. When discussing engineering, the system level and the stage in the plant construction value chain determine the optimal approach.

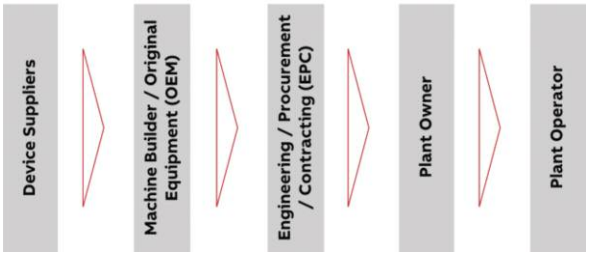


Figure 2: Plant construction value chain.

Plant lifecycle

The value chain described above explains how a plant is planned and built, but once it exists, it must be operated and maintained. Even in operation, the products that are being produced are engineered, sometimes to produce new product families, sometimes to produce according to customer specifications. In the plant value chain, it may be necessary to design some equipment specifically for a particular purpose (in size, performance, or special industry requirements). Changes in the product engineering may then result in a re-engineering of the plant, i.e. in a change of its production sequence or machining instructions.

But even the most reliable components may fail from time to time. Without proper maintenance, they fail even earlier.

The maintenance tasks themselves may require some engineering, since a component may no longer be available, or a newer, better component may require an adaptation of the production process as well, i.e. it requires re-engineering.

Technology challenges

The engineering processes discussed above pose a few technological challenges. Industrial applications typically require high capital investment for the equipment, and many processes handle high energy (electrical, thermal, mechanical, etc.). In short: industrial applications are expensive and dangerous. Failures not only reduce the production of the plant through outages, but they can cause injuries and fatalities, or even major environmental damage with long-lasting consequences.

Industrial AI

When looking into machine learning in engineering, we first need to consider a few boundary conditions that are imposed by the nature of industrial applications.

Safety

Industrial applications are very often dangerous. Safety regulations require the equipment to meet reliability requirements in order to be applied in certain industries. Today's neural networks, however, cannot guarantee a reliable reaction in a given situation. At times even a well-trained system responds incorrectly. In situations where the AI system supports the human, any correct conclusion is adding to the human's performance. False negatives may still be caught by the human, but if not, the performance is not worse than the human's. However, false positives reduce the trust in the system and should be avoided. In a safety environment, false negatives are not acceptable. Any situation not properly caught may lead to damage. False positives, however, are accepted: an industrial safety system shuts down in order to absolutely not miss a safety incident. This also relates to engineering. A wrong decision in the engineering process may reduce the reliability of the equipment in the field, without being discovered. Hopefully this would get detected in acceptance tests, but by that stage it is already produced, and the mistake results in delays and quality costs.

Data

Compared to consumer data, industrial data is sparse. There are probably more industrial machines than consumers, but data (measurements or engineering data) of an electrical breaker cannot be compared to data from a motor or a welding robot. The data-hungry ML algorithms that are fed with millions of images or text samples starve on the data available from industry. Compared to real-time plant measurements, engineering data is even more scarce. Some equipment is designed and produced once for one individual customer. Good industrial use cases may produce thousands or even millions of products, but they are only engineered once. ML in engineering therefore is challenging to train using consumer ML methods. Furthermore, engineering data contains all the intellectual property that went into designing a product. Industrial customers are very reluctant to share measurement data from the factory floor and engineering data is shared even less. To train an ML system compiling similar engineering data from a variety of companies is even less desirable.

Simulation & digital twin

A part of the data challenge can be addressed by using simulation to generate the data. Like reinforcement learning algorithms that played Go until they mastered it, a simulator can provide the data to learn how to build a product. But while Go is played along a limited set of rules on a limited board, a design simulator (product or plant level) follows the more complex laws of physics. The complexity of such an approach is much higher than winning a board game.

Given the challenge of getting enough engineering data from real plants, simulation is probably one of the few promising approaches to solve the engineering challenge using ML. However, a good simulator that properly reflects reality is expensive and not easy to build. Furthermore, a simulator that properly reflects the true plant's behavior, i.e. its behavior mapped to measured values, is hard to achieve: mapping the parameters is difficult, but mapping them repeatedly over time, factoring in wear and ageing, makes this task a continuous challenges. To map those aspects that are difficult to model through physics, adding ML to map the difference between the 'as built' behavior and the operation over time is one approach that has been tried successfully. Once a good simulator is available, the choice between a good optimiser and an ML system needs to be considered.

In today's digitalised industrial environment, simulation is seen as an important component of a digital twin. The digital variant of a product or system, that behaves exactly like the real plant, has its benefits in designing the plant, but when extended with other digital aspects of a component, e.g. real-time measurements, or maintenance records, can serve many more business models along the plant lifecycle. In the context of a digital twin, simulation capabilities may therefore be worth developing beyond the ML-based engineering stage and may make a business case valuable.

Autonomous industrial systems

As we discussed engineering, it reflects the approach to design, build and operate equipment, e.g. a plant. Changes in the plant's purpose, e.g. to build a product variant or optimize production, requires re-engineering of part of the process, and it's adaptation on the plant floor. An alternate approach could be envisioned: a plant that reacts to a change of its purpose, and autonomously finds a way to produce what it is asked for. Such a plant requires capabilities that go beyond current engineered systems. To pre-engineer all possible variants into the system and program it's automation system accordingly exceeds the scope of today's automation system in size as well as complexity.

Such a system would need to have a basic understanding of its own capabilities, and the proper knowledge of how its capabilities could be employed to produce what it is asked for. Engineering of such a system is then an inherent, AI-supported capability of the system itself. Such systems redefine engineering and operation, and provide a flexibility in production that is today unseen.

Please contact:
Christopher Ganz
Head of Solutions & Standards ABB Future Labs,
christopher.ganz@ch.abb.com

Enhancing Technical Simulations with Machine Learning

by Hamid Asgari, Juha Kortelainen, and Mikko Tahkola (VTT)

Artificial intelligence, machine learning and artificial neural networks are introducing interesting opportunities to engineering design as well as to monitoring and operations of systems and processes. Core and enabling technologies are evolving fast, with great potential for industrial applications. Machine learning, combined with simulation, can enable models that are both fast and sufficiently accurate to be used in new applications.

Modelling and simulation are commonly used tools in mainstream engineering design. For instance, numerous open source and propriety software applications are available for structural analysis, computational fluid dynamics, and system simulation. These engineering tools have revolutionised the design process, enabling increasingly demanding designs to be completed more quickly and efficiently. In engineering design use, the performance of modelling and simulation tools is becoming increasingly important. The use of large design studies and application of mathematical optimisation are emphasising the computational performance of simulations, especially when thousands of simulations are needed. Increased performance of each simulation can remarkably improve the overall analysis time and enable new information to be gained faster. This means that larger, more thorough studies can be completed within a given time.

The current interest in digital twins has also emphasised the need for faster simulations and improved computational efficiency. A digital twin is a digital, i.e., computer-based, replica of a real-world product, system or process. The digital twin represents the relevant features (from a use and operation perspective) of the real-world twin and synchronises its state with it. One way to define and categorise digital twins is to contemplate their ability to represent the state of the real-world system in time. If the digital twin is able to describe the history and present state of the real twin, but cannot predict its future state, the digital twin does not necessarily need any simulation features and can be considered a "descriptive digital twin". On the other hand, if the digital twin can predict the future state of the real twin, some simulation capabilities are usually needed and the digital twin can be categorised as a "predictive digital twin". Furthermore, if the digital twin is used to optimise the function and operation of the real twin, some mathematical optimisation capabilities are generally needed, and we can categorise it as a "prescriptive digital twin". The more prescriptive the use of a digital twin, the more its computational efficiency matters. Thus, innovative ways to simulate complex systems, especially involving complex physical phenomena, are needed.

A surrogate model is a simplified and usually computationally efficient replacement of the original or more accurate model of the target. The original model can be based on physics-based simulation, which can require considerable computing resources and be very time consuming. A solution may lie in machine learning and artificial neural networks, which are currently being researched and developed to produce data-driven models.

Machine learning is a subset of artificial intelligence that deals with exploring the data structure and fitting the data into a model. It relies on the use of computers, algorithms, and data processing techniques for clustering, regression, classification, and pattern recognition. Advances in artificial intelligence and machine learning have helped improve engineering techniques in different disciplines in ways that may not have been possible with conventional methods. Machine learning algorithms are grouped into three main categories: supervised learning, unsupervised learning, and reinforcement learning. Methods utilising so-called deep artificial neural networks are commonly called deep learning. Deep learning is employed to set up a complex artificial neural network structure with multiple layers to train big datasets with complex interconnection relationships. Figure 1 illustrates a classification of machine learning algorithms and some of their applications [L1].]. Figure 2 represents applications of machine learning in industry for fault detection, prediction and prevention [1]. As Figures 1 and 2 show,

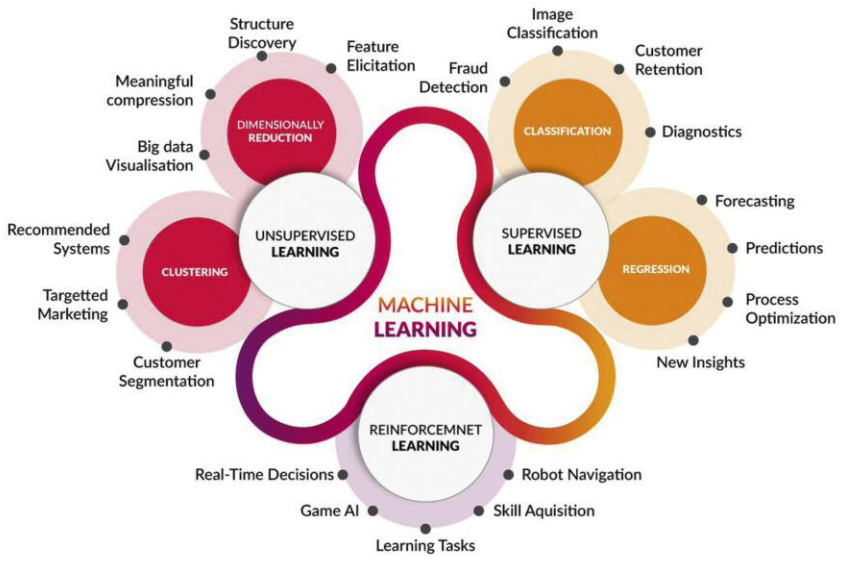


Figure 1: A classification of machine learning algorithms [1].

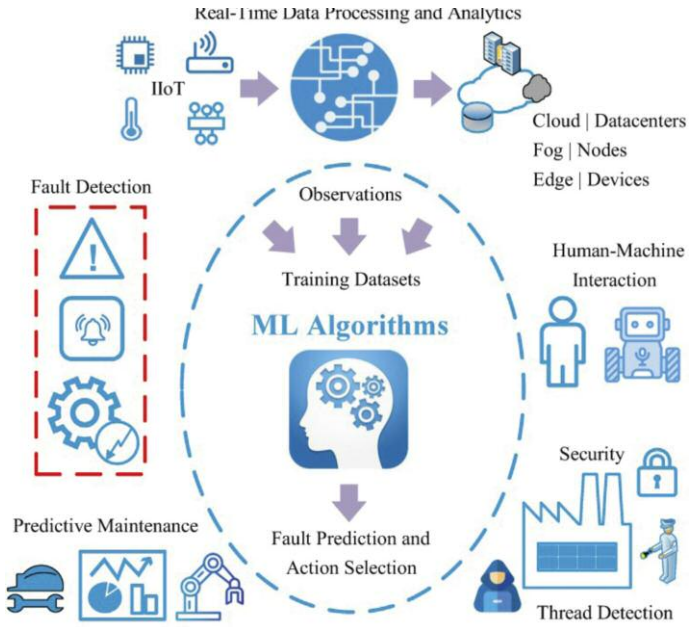


Figure 2: Applications of ML in industry for fault detection, prediction and prevention [2].

machine learning may be employed in a variety of engineering applications, such as process optimisation, predictions, diagnostics, big data visualisation, and robot navigation.

Artificial neural networks have been successfully used in various applications across different domains and can be applied to surrogate modelling as well. A range of open source software for building artificial neural network models is available. The surrogate approach is model agnostic in the sense that only the data from the simulation software is required. Depending on the simulation speed, it can be useful to be able to parallelise the data generation to gather the data faster. Similarly, it is possible to speed up surrogate model development by parallelising the artificial neural network training process.

We investigated the use of artificial neural networks to create a surrogate model of a physics-based industrial process simulation model. Modelled systems were a liquid level-controlled water tank process, and a methanation reactor in a power-to-gas process, shown in Figure 3 [2]. In the former case, normalised root mean squared error of the surrogate model was 0.01% and in the latter about 4%. In the electrical machine domain, we have created surrogate models of a permanent magnet synchronous motor that are about 2,500 times faster than the original model, with a normalised root mean squared error of between 0.5 and 10%, depending on the operating point. The case studies show that surrogate modelling is a potential tool to enhance research, development and design work efficiency by enabling faster simulation.

Link:
[L1]: https://kwz.me/h0o

References:
[1] A. Angelopoulos, et al.: "Tackling Faults in the Industry 4.0 Era—A Survey of Machine-Learning Solutions and Key Aspects", Sensors, Basel, 20(1): 109, 2020.
[2] M. Tahkola: "Developing dynamic machine learning surrogate models of physics-based industrial process simulation models", master's thesis, University of Oulu, 2020.

Please contact:
Hamid Asgari, Juha Kortelainen, Mikko Tahkola, VTT, Finland
ext-hamid.asgari@vtt.fi,
juha.kortelainen@vtt.fi,
mikko.tahkola@vtt.fi

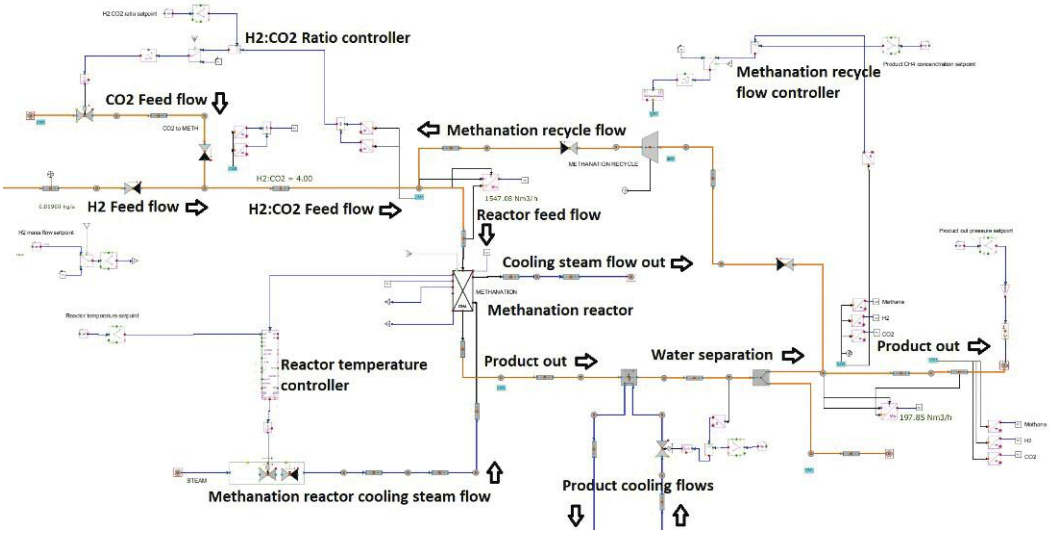


Figure 3: Physics-based simulation model of the methanation reactor in a power-to-gas process [3].

Guaranteeing Performance Specifications for Vehicle Systems with Learning Agents through the Robust Control Theory

by Balázs Németh and Péter Gáspár (SZTAKI Institute for Computer Science and Control)

Several advanced complex control systems can incorporate learning agents, especially in designing functions in automated vehicles. At the same time, an open problem is to find a systematic design method that is guaranteed to satisfy the performance specifications of the system. This paper presents a possible design method based on robust control theory, which has been developed in the Systems and Control Laboratory of SZTAKI.

Highly complex automated vehicle and transportation technologies are evolving as a means of addressing the problems of traffic congestion, energy consumption and emissions. The increasing complexity of control and decision tasks has resulted in the combined application of various control systems, e.g. model-based robust and optimal control, nonlinear control and machine learning-based solutions. One of the most important fields is related to the control of automated and autonomous vehicles, in which several driving features must be automated to reduce the role of human intervention, e.g. sensing the environment, making decisions, trajectory design, control and intervention with smart actuators [1].

The increasing complexity of the control systems poses the challenge of applying control methods that will guarantee performance specifications. Road stability, manoeuvrability and safety are primary performance requirements in safety-critical systems, which must always be guaranteed by the control during the operation of the system. Other variables, including comfort, energy consumption and emissions reduction are considered "secondary performances", which must be considered by the control system, but may be violated in critical situations, e.g., if a vehicle collision or pedestrian accident is predicted.

The complexity of the control requires novel analysis, synthesis and validation methods that can guarantee the primary performances and possibly the secondary performances as well. Model-based control design methods have advantages in terms of theoretical guarantees for the performances, but the high complexity of the control-oriented model and the large number of performance specifications in the control design must be limited for numerical reasons in the mathematical computation of the control and the practical implementation possibilities. However, in the case of learning-based techniques the control system may have high complexity, e.g., convolutional neural networks, and there are effective methods for the learning solutions. Although the different types of enhanced learning control methods can solve various control tasks effectively, their achieved performance level is not theoretically guaranteed. The quantity and quality of the learning samples can be selected at any size, but this does not guarantee the avoidance of performance degradation in an emergency scenario or robustness against faults and disturbances. The validation of learning-based automated vehicle control systems poses similar challenges. There exist conventional test scenarios for the validation of the model-based controllers in driver assistance systems, but the evaluation of control systems in automated vehicles may require huge number of scenarios. The problem is to find a theoretically emphasized process for the limitation on the test scenarios, with which the guarantees on performance requirements can be evaluated.

Although learning control can provide partial theoretical results, a general systematic solution does not exist. We aim to provide a design framework based on the robust Linear ParameterVarying (LPV) analysis and synthesis, in addition to learning methods, to guarantee performances [2].

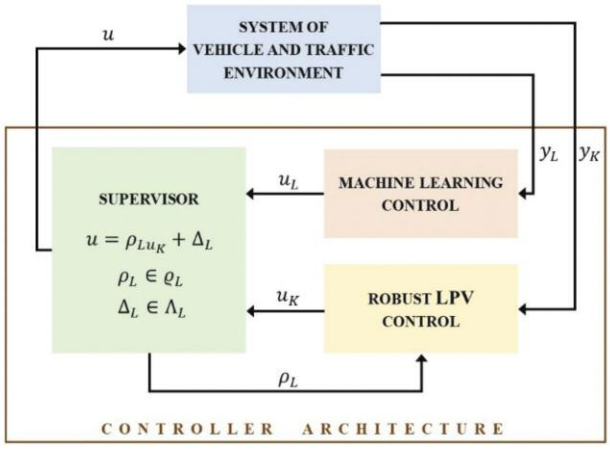


Figure 1: Scheme of the design framework with the control components and the supervisor.

Design framework to achieve guaranteed performances

In the structure of the control system a learning control and a robust LPV control operate together under the monitoring of a supervisor, see Figure 1. The learning control is designed to consider the specifications on primary and secondary performances through its agents. It uses various information sources as measured signals in y_L. The robust LPV control guarantees the primary performances, while several secondary performances might not be considered in its design. It uses only onboard signals of the automated vehicle, y_K. Its scheduling variable ρ_L comes from the supervisor, which applies both con-

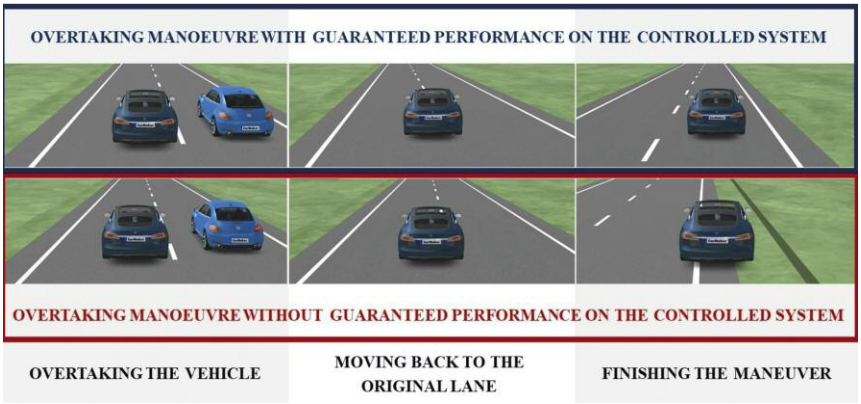


Figure 2: Illustration of the overtaking maneuvre in connection with the guaranteed performance.

trollers and monitors both to vehicle motions and the traffic environment. In general, especially under normal travelling conditions, the supervisor uses a control signal which is calculated by machine learning control. However, dangerous situations may occur in which guaranteeing primary performances is essential. In these cases, the supervisor uses the control signal of the robust controller, overriding the current control signal.

The role of the supervisor is to monitor the signals of the control elements and to make a decision about the intervention. The rule of differentiation between nominal and critical scenarios is represented by an optimization process in the supervisor, which is based on its input signals u_L and u_K as follows. The control signal u is formed as $u=\rho_L u_K+\Delta_L$, where ρ_L elements of ϱ_L and Δ_L elements of Λ_L are ρ_L, Δ_L are scalar values and ϱ_L, Λ_L represent their domains. Since ϱ_L, Λ_L are considered in the design of the robust LPV controller as scheduling variable and known uncertainty, u in the environment of u_K guarantees the minimum level of the primary performances. Thus, if u_L is inside of the environment of u_K, $u=u_L$ is selected. But, if u_L is outside of the environment of u_K, then $u\neq u_L$ and u is computed through the saturation of u_L. Thus, the control signal achieves at least the minimum performance level on the primary performances in both cases and, moreover, the secondary performances can also generally be achieved.

The rationale behind the application of the robust LPV formalism is that this method is well elaborated and is used to address various industrial problems [2]. The advantage of the proposed solution is that it is independent of the internal structure of the learning control methods, i.e., it can be used for deep learning, reinforcement learning and other methods. Another advantage is that the negative impacts of the degradation in the external information sources (e.g. loss of internet communication) on the performances can be avoided, since y_K contains only onboard measurements.

Applications within automated vehicle control systems

The proposed method has been applied at various levels of automated vehicle control tasks in simulation environments. On the level of local control, the method is applied to the design of steering control. Deep-learning-based end-to-end learning solutions have high impact on automated driving, which uses visual signals for the actuation of the steering system. In practice, the learning can be performed through reinforcement learning or machine learning through the samples of a vehicle driven by an expert driver. As a result, steering control can provide an acceptable path following functionality and traveling comfort, but unfortunately, the performances are not guaranteed theoretically. The robust LPV control can be designed based on a simplified physical vehicle model, in which the path following can be theoretically guaranteed.

On the level of vehicle functionalities, the proposed framework has been applied to cruise control design. Several performances must be involved in the control design, e.g. travelling time, speed limits, vehicle tracking, energy consumption. The control solution can be achieved with complex nonlinear optimization algorithms, but its implementation has numerical limitations. The motivation of learning control is to reduce online computation through neural networks which are trained through supervised learning methods on the offline solutions of the optimization problem. Since keeping speed limits and distances from the surrounding vehicles are safety performances, they can be guaranteed by the robust LPV control, which is designed based on a simplified vehicle model.

On the level of multi-vehicle interactions, the method has been applied to the solution of overtaking tasks of automated vehicles. The role of the overtaking strategy is to find a trajectory for the automated vehicle with which safe motion can be guaranteed. It requires information about the objects in the vehicle's environment, from which the classification of the objects and their motion prediction can be performed. The learning process has been performed through various multi-vehicle scenarios, whose results are a route and a speed profile. Despite the large training set, there may be multi-vehicle scenarios which result in inappropriate vehicle motion. Figure 2 illustrates an example in which an automated vehicle overtakes a slower vehicle [3]. The highlighted scenarios are related to the same time. The control strategy with guaranteed performance ensures a safe completion of the overtaking manoeuvre, even if the machine learning control provides unacceptable vehicle motion.

References:

[1] P. Gáspár, B. Németh: "Predictive Cruise Control for Road Vehicles Using Road and Traffic Information", Springer Verlag, 2019.
[2] P. Gáspár, et al.: "Robust Control Design for Active Driver Assistance Systems", Springer Verlag, 2017.
[3] B. Németh, T. Hegedűs, P. Gáspár: "Performance Guarantees on Machine-Learning-based Overtaking Strategies for Autonomous Vehicles", European Control Conference, pp. 136-141. 2020.

Please contact:

Balázs Németh, Péter Gáspár
SZTAKI, Hungary
balazs.nemeth@sztaki.hu,
gaspar.peter@sztaki.hu

Machine Learning for Aerodynamic Uncertainty Quantification

by Dishi Liu, Daigo Maruyama and Stefan Görtz (German Aerospace Center)

Within the framework of the project "Uncertainty Management for Robust Industrial Design in Aeronautics" (UMRIDA), funded by the European Union, several machine learning-based predictive models were compared in terms of their efficiency in estimating statistics of aerodynamic performance of aerofoils. The results show that the models based on both samples and gradients achieve better accuracy than those based solely on samples at the same computational costs.

The UMRIDA project (Uncertainty Management for Robust Industrial Design in Aeronautics) is a collaborative project funded by the European Union that aims at a technology readiness level of robust design under a large number of simultaneous uncertainties. This work on machine learning-based predictive models was done in the UMRIDA project for a more efficient quantification of uncertainties in aerodynamic performance, by the Institute of Aerodynamics and Flow Technology in German Aerospace Center at Braunschweig, Germany.

In the context of uncertainty quantification and robust design, the typical quantities of interest are statistics of drag coefficient (C_d) and lift coefficient (C_l) of aerofoils. Uncertain input parameters are operational parameters such as the angle of attack, the Mach number, the Reynolds number, and an inherently large number of geometric parameters.

We made two comparisons of various methods in their efficiency of quantifying aerodynamic performance uncertainties caused by operational and/or geometric uncertainties. First, we compare two surrogate integration methods based on machine learning processes, Gaussian process model (Kriging) and gradient-enhanced Kriging (GEK), with a direct integration method based on quasi-Monte Carlo (QMC) quadrature on a viscous test case. Second, we compare the QMC quadrature and four machine learning-based integration methods, namely GEK, polynomial chaos combined with a sparse Gauss-Hermite quadrature, gradient-enhanced radial basis functions (GERBF) and a gradient-enhanced polynomial chaos collocating method on an inviscid test case.

Viscous test case: Comparison of Kriging and GEK

This comparison is based on a CFD model of the viscous flow around the RAE2822 aerofoil. We opt for DLR's unstructured RANS solver TAU [1], the Spalart-Allmaras turbulence model, and a central flux discretisation scheme. The domain is discretised by a hybrid unstructured grid in which the aerofoil has 380 surface nodes. The uncertainties come from a random Mach number and angle of attack, together with a random perturbation to the original aerofoil geometry at every surface grid point. The two operational variables are assumed to be beta-distributed around M=0.729 and α=2.31°, respectively. The perturbations in Mach number and angle of attack are with a support within ±2% of the nominal values. The geometry perturbation is modelled by a random field parameterised into 24 independent Gaussian variables through a truncated KLE [2].

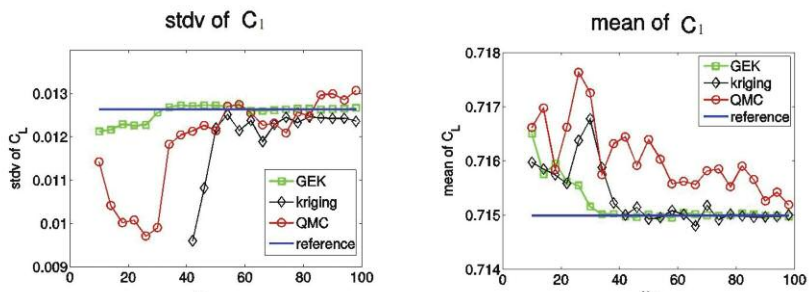


Figure 1: Convergence of estimates of Cl statistics to the reference statistics, reproduced from [3] with permission.

Quasi-Monte Carlo quadrature and two machine learning-based UQ methods, Kriging and gradient-enhanced Kriging, are applied to the test case and their efficiency is compared in estimating two statistics (mean, standard deviation) of the coefficient of lift (Cl). The accuracy of the estimates is judged by comparing to reference statistics which are based on 10,000 QMC samples. The computational cost is measured in terms of "compensated evaluation number" Nc to take the cost of gradient evaluation into account.

Figure 1 shows the error convergence in the two statistics of Cl by using various methods. GEK is seen to converge faster than Kriging in estimating both statistics. This can obviously be attributed to the greater amount of information utilised by the former at the same computational cost with the help of the adjoint TAU solver. It can also be observed that if the sample number is small Kriging may perform worse than QMC.

Inviscid test case: Comparison of four machine learning-based integration methods and direct integration

This comparison is based on a CFD model of the inviscid flow around the RAE2822 aerofoil at a Mach number of 0.73 and a 2.0° degree angle of attack. We use the TAU flow solver, opting for a central flux discretisation, scalar dissipation, and a backward Euler solver. The domain is discretised by a 193-by-33 structured grid in which the aerofoil has 128 surface nodes.

The source of uncertainty is a random perturbation to the original aerofoil geometry, which is modelled by a random field parameterised into nine independent Gaussian variables through a KLE [2]. The five methods in the comparison are applied to the test case and compared in terms of their efficiency in estimating statistics of Cl and

Cd, as well as the probability distribution functions (pdf). The reference values of these statistics are obtained from an integration of four million quasi-Monte Carlo (QMC) samples of the CFD model.

The results of this comparison are detailed in [3]. It is observed that generally the gradient-employing machine learning-based methods perform better than direct integration methods. This can be ascribed to the fact that the former utilises more information at the same computational cost. This advantage comes from the cheaper cost of the gradients computed by an adjoint solver in the case that the number of response quantities of interest is smaller than the number of variables (in our case, two versus nine), and the advantage would increase for a larger number of variables or fewer response quantities. The results indicate that GEK and GERBF are the most efficient methods. Besides the cheaper gradients, this could be attributed to properties of the kernel functions they use and the effort of tuning model parameters.

References:

[1] T. Gerhold, V. Hannemann, D. Schwamborn: "On the validation of the DLR-TAU code", in New Results in Numerical and Experimental Fluid Mechanics, in: Notes on Numerical Fluid Mechanics, vol. 72, Vieweg, pp. 426–433, 1999.

[2] D. Liu, A. Litvinenko, C. Schillings, V. Schulz: "Quantification of airfoil geometry induced aerodynamic uncertainties - comparison of approaches", SIAM-ASA journal on Uncertainty Quantification, vol. 5, no. 1, pp. 334–352, 2017.

[3] D. Maruyama, D. Liu, and S. Görtz: "Comparing surrogates for estimating aerodynamic uncertainties of airfoils", in Uncertainty Management for Robust Industrial Design in Aeronautics. Springer, 2019, pp. 213–228.

Please contact:
Dishi Liu
TU Braunschweig, Germany
dishi.liu@tu-bs.de

Machine-Learning-Based Reduced Order Model for Macro-Scale Stochastic Plasticity

by Emir Karavelić (Univ. of Sarajevo), Hermann G. Matthies (TU Braunschweig) and Adnan Ibrahimbegovic (Univ. de Technologie de Compiègne)

Bayesian inference can be used in machine learning to provide a reduced order model for multi-scale stochastic plasticity, with parameters as random variables. Machine learning can deliver either the random variables or their probability measure.

We are developing a machine learning procedure that provides a probability-based scale bridging for concrete composite material. This method is appropriate for situations in which the scale separation of standard homogenisation approach does not apply. The proposed procedure is capable of passing the detailed information available at the micro-scale (at which different phases are visible) to the chosen reduced order model at the macro-scale (where concrete is a single continuum). This is accomplished, not only for representing elastic, but also different phases of inelastic response, resulting in a stochastic plasticity model for localised failure. This proposed machine learning procedure exploits Bayesian inference to provide the probability distributions of such stochastic plasticity model parameters expressed as random variables (RV). Two different approaches of Bayesian inference can be used to quantify the uncertainty: one constructs the RV, the other the probability measure.

This work contributes to machine learning formalism through the development of stochastic constitutive models that capture localised failure sensitivity to initial and induced defects for structures built of heterogeneous composite material, such as concrete. Such models can enable better predictions of crack spacing and opening, which can be used to improve concrete durability. This information is relevant to many industrial applications, given that concrete is probably the most widely used material in construction and that many existing concrete structures are rapidly ageing.

When it comes to constructing predictive stochastic plasticity fracture models, this kind of composite material has favourable features: two-phase microstructure with aggregate vs. cement paste visible at micro-scale [1], non-local dimension owing to typical fracture modes of massive structures with significant contribution of the fracture process zone (FPZ), and a fabrication that is comparable to large-scale additive-manufacturing, where the complete structure is cast with the same material rather than an assembly of various components. The latter is the crucial hypothesis to make the proposed macro-scale stochastic plasticity model feasible in terms of predicting probability-distribution-based estimates of structure properties. The main novelty here is that a standard homogenisation approach no longer applies and all scales are to be treated probabilistically. Thus, our proposed approach is to capture not only the average response on the larger scale, but also the potential variability bounds. This is particularly important for the testing of heterogeneous cement-based composites, where the scales in the test specimen are generally poorly separated. In this work we focus in particular on the uncertainty propagation which allows the inelasticity to be connected at multiple scales, starting from the fine scale (here the micro-scale for concrete) where the heterogeneous composite failure mechanisms and the variability of model parameters

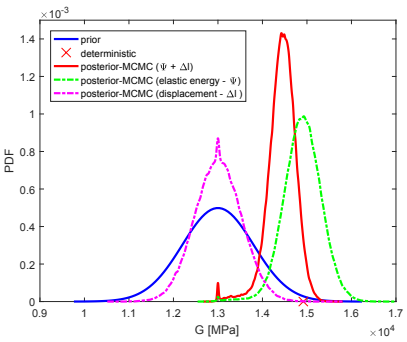


Figure 1: Comparison of prior and posterior pdf functions for shear modulus performed on uniaxial compression test obtained by MCMC (one realisation on micro-scale – w_1).

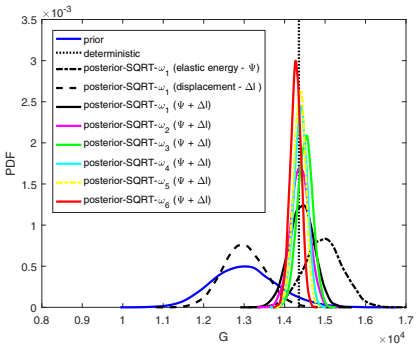


Figure 2: Comparison of prior and posterior pdf functions for shear modulus bulk modulus performed on uniaxial compression test obtained by SQRT Kalman filter (six realisations on micro-scale – w_6).

can be captured much better by a corresponding representation of different phases (here aggregate vs. cement paste).

Our main focus in this work is on the model reduction from micro-scale to macro-scale, which defines generalised an ED-FEM in a probability framework and allows for stochastic coarse graining. This is illustrated in switching from micro-scale to macro-scale, to provide a replacement with a generalised ED-FEM once the crack pattern inside the corresponding micro-scale element is stabilised. Namely, the goal is to then replace the micro-scale computation with the corresponding stochastic macro-scale model of such an ED-FEM. We illustrate this idea on a plasticity model, where such parameters also include yield stress (for defining the fracture process zone — FPZ) and ultimate stress (for defining localised failure).

Two different methods for Bayesian inference have been tested and compared in the proposed approach, both based on the Bayes theorem that makes it possible to incorporate new information that has been generated in a particular loading program. Each unknown parameter of the reduced model is modelled as a random variable. Such a description has two constituents, the measurable function representing the random variable RV, and the measure. Markov chain Monte Carlo (MCMC) updates the measure, whereas the various filters [2] change resp. update the measurable function (Figures 1 and 2). We formulate both methods as functional spectral approximations of stochastic problems and introduce, in combination with the second method, a new procedure that does not need any sampling, hence works with the subsequent update in a deterministic manner. It also seems to be the fastest and most reliable method compared with others. We show by example that the spectral resp. polynomial chaos version of the Gauss-Markov-Kalmán filter (GMKF) also works for highly nonlinear non-smooth problems with non-Gaussian measures (Figure 3 and 4). More detailed information can be found in [1,2,3].

References:

[1] E. Karavelić, et al.: "Concrete micro-scale model with full set of 3D failure modes with random distribution of aggregate and cement phase", Part I: Formulation and numerical implementation, Comp. Methods Appl. Mech. Eng., 344, 1051–1072, 2019.

[2] H. G. Matthies, A. Ibrahimbegovic: "Stochastic Multiscale Coupling of Inelastic Processes in Solid Mechanics", in: M. Papadrakakis, G. Stefanou (eds.), Multiscale Modelling and Uncertainty Quantification of Materials and Structures, 3, 135–157, (2014), Springer. doi: 10.1007/978-3-319-06331-7-9

[3] A. Ibrahimbegovic, H. G. Matthies, E. Karavelić: "Concrete micro-scale model with full set of 3D failure modes with random distribution of aggregate and cement phase", Part II: reduced model of macro-scale stochastic plasticity identification by Bayesian inference, Comp. Methods Appl. Mech. Eng., in press, 2020.

Please contact:

Adnan Ibrahimbegovic
Universite de Technologie de Compiègne, France
adnan.ibrahimbegovic@utc.fr

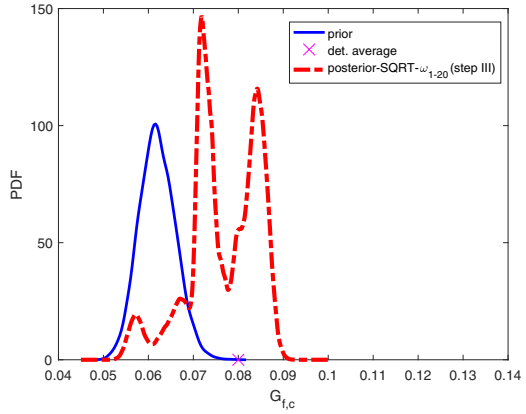


Figure 3: Updates for softening parameter $G_{f,c}$ (fracture energy) obtained by SQRT Kalman filter with three measurements taken from uniaxial compression test for 20 different realisations – w_{20}

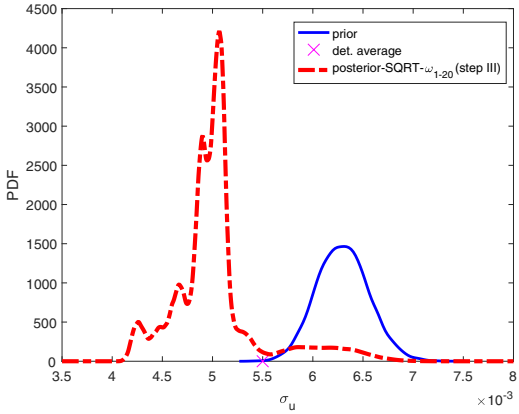


Figure 4: Updates for softening parameter s_u (limit stress) obtained by SQRT Kalman filter with three measurements taken from uniaxial compression test for 20 different realisations – w_{20}

Deep Neural Network-Based Filtering Techniques for Data Assimilation

by Truong-Vinh Hoang (RWTH-Aachen University) and Hermann G. Matthies (TU Braunschweig)

Data assimilation is a challenge in many forecasting applications ranging from weather and environmental forecasting to engineering applications such as structural health monitoring and digital twins. A common technique for data assimilation is the ensemble Kalman filter (EnKF). However, it is well-known that the EnKF does not yield consistent estimates for highly nonlinear dynamical systems. Deep learning (DL) techniques can be applied to improve the EnKF in high-dimensional and nonlinear dynamical systems. This article presents an extension of the EnKF using deep neural networks (DNNs) with a focus on the theoretical and numerical aspects.

Data assimilation aims to update states of a dynamical system by combining numerical models and observed data which can be sparse in space and time. Owing to uncertainty in the numerical models as well as measurement data in the probability setting, knowledge about model states is presented using probability distributions. When new measurement data become available, the knowledge about the model states is updated by conditioning the state distribution on the measured observations, which is usually performed using Bayes' theorem. For high-dimensional and nonlinear simulation models, an accurate representation of the assimilated state distribution comes at an extremely high computational cost. A common method for data assimilation with an acceptable computational budget is the EnKF. In this method, the state distribution is approximated by an ensemble, and the assimilation is performed by applying the Kalman filter on each ensemble's member. It is well-known that the EnKF is not appropriate for highly nonlinear dynamical systems due to linear approximations of the dynamical systems and observation maps in the Kalman filter. Thus, there is a need to develop ensemble filtering methods that perform better in these situations.

In general, a filter is a function of the observations and the current states mapping to the assimilated states. This map can be very complex, especially for high-dimensional state spaces. Deep learning, which has significant advantages in representing complex functions between high-dimensional spaces, has great potential to be applied in these problems. Indeed, the general idea here is to use DNNs to construct filtering maps such that the assimilated ensemble approximates the conditioned distribution yielded by Bayes' theorem. The datasets using for training the DNNs are the ensembles of states and predicted observations.

In particular, we are developing a novel DL-based ensemble conditional mean filter (EnCMF). The EnCMF is a generalisation of EnKF for nonlinear dynamical systems with non-Gaussian distributions of states and measurement errors [1]. An implementation of the EnCMF has been developed using polynomial chaos expansions for approximating the conditional expectation (CE) of mean values. However, this approximation is not suitable for high-dimensional state spaces due to the curse of dimensionality, i.e., the number of polynomials in the expansion increases exponentially with respect to dimensionality. In the DL-based EnCMF, we approximaIn the DL-based EnCMF, see Figure 1 for its implementation procedure, we approximate the CE of mean values by a DNN.te the CE of mean values by a DNN. Thanks to the orthogonal property of CE, the loss function used to train this DNN is the mean squared error—a commonly used loss criterion for DNNs. The trained DNN is then used to form the filter.

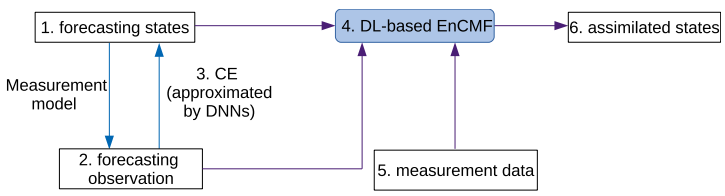


Figure 1: Implementation procedure of the DL-based EnCMF with 6 steps: (1) evaluating an ensemble of the forecasting states, (2) evaluating the corresponding ensemble of the forecasting observation, (3) approximating the CE of the mean using a DNN, (4) constructing the DL-based EnCMF, and (5, 6) plugging the measurement data into the DL-based EnCMF, and computing the assimilated ensemble.

Unlike the EnKF, the DL-based EnCMF does not linearise the dynamical system and observation maps. In comparison with the EnKF, the DL-based EnCMF yields better estimates, but it requires larger ensemble sizes. For example, by increasing the size of ensembles, the mean value of the state ensemble converges to the conditioned mean yielded by the Bayesian formulation—a property that cannot be obtained using the EnKF. A numerical challenge of the DL-EnCMF is the limit size of data sets—the ensembles of states and predicted observations—which can lead to the over-fitting problem when training DNNs. A way to ease the over-fitting phenomenon is to use techniques such as regularisation, dataset augmentation and noise robustness when training the networks.

In the future, we will investigate other DL-based filters, e.g., using conditional expectations of higher-order moments or the variational Bayesian inference. Moreover, training algorithms such as those combining online-offline training sessions to reduce the online training computational cost will be considered.

Reference:

[1] H. G. Matthies et al.: "Parameter estimation via conditional expectation: a Bayesian inversion", Adv. Model. and Simul. in Eng. Sci. (2016) 3:24. DOI: 10.1186/s40323-016-0075-7

Please contact:

Truong Vinh Hoang,
RWTH-Aachen, Germany
hoang@uq.rwth-aachen.de

Herman G. Matthies, Technische Universität Braunschweig, Germany
h.matthies@tu-braunschweig.de

Using Deep Learning and Data Integration for Accurate Rainfall Estimates

by Gianluigi Folino, Massimo Guarascio (ICAR-CNR), Francesco Chiaravalloti and Salvatore Gabriele (IRPI-CNR)

Accurate rainfall estimates are critical for areas presenting high hydrological risks. We have devised a general machine learning framework based on a deep learning architecture, which also integrates information derived from remote sensing measurements, such as weather radars and satellites. Experimental results conducted on real data from a southern region in Italy, provided by the Department of Civil Protection (DCP), show significant improvements compared to current state-of-the-art methods.

Accurate rainfall estimates are important in a range of fields, including meteorology, geology and agronomy, allowing researchers to model hydrological and other environmental processes. The spatial variability of rainfall greatly affects the local hydrological processes, and the prediction accuracy of rainfall-runoff simulations is strongly determined by the precision of rainfall estimates [1]. Therefore, an accurate retrieval of the spatio-temporal rainfall patterns is crucial for flood hazard protection, river basin management, erosion modelling and other applications for hydrological impact modelling.

To this end, rainfall sensors, called "rain gauges", are commonly used by meteorologists and hydrologists to obtain direct, quantitative, and reliable measurements of rainfall intensity in single point sites. Spatial interpolation methods can use rain gauge data to extract an estimate of the precipitation field over a broader area. However, even dense rain gauge networks may be too sparse and insufficient to reconstruct the rainfall field, failing to capture heavy convective events.

Recent research has investigated the possibility of integrating data from heterogenous sources to improve the accuracy of rainfall estimation models. However, traditional interpolation methods (e.g., ordinary kriging) only handle a single source at a time. Kriging with external drift (KED) was introduced to overcome this, but its high computational cost makes it difficult to use in a real-time setting [2].

In this scenario, deep learning-based architectures have the potential to efficiently extract accurate rainfall estimation models by combining raw low-level data recorded by heterogenous data sources. Indeed, we exploit the capacity of deep neural networks (DNNs) to work in a hierarchical way: i.e., several layers of non-linear processing units are stacked into a hierarchical scheme and each subsequent layer generates a feature set with a higher level of abstraction than the previous one. Therefore, deep learning-based approaches are the ideal choice to analyse raw data provided in different formats and from different types of source. Moreover, deep architectures can be used within infrastructures for big data storage and analysis (e.g., Hadoop and Spark) and can exploit GPUs to parallelise the computation and to reduce the learning times.

A joint collaboration between ICAR-CNR and IRPI-CNR aimed to design a framework [3] based on three main macro-components (discussed in further detail below): (i) information retrieval, (ii) data analytics and (iii) evaluation, making it possible to integrate information extracted from many data sources.

The information retrieval macro-module is designed to extract and integrate data from different sources. Specifically, a "data source connector" is used to establish the connection with a specific data source. The information extracted from each connector is provided as input for the "data wrapper" module that combines these data into a single view suitable for the analysis. Finally, the raw data is stored in the "knowledge base" (KB), which is used for data exchange among the framework modules.

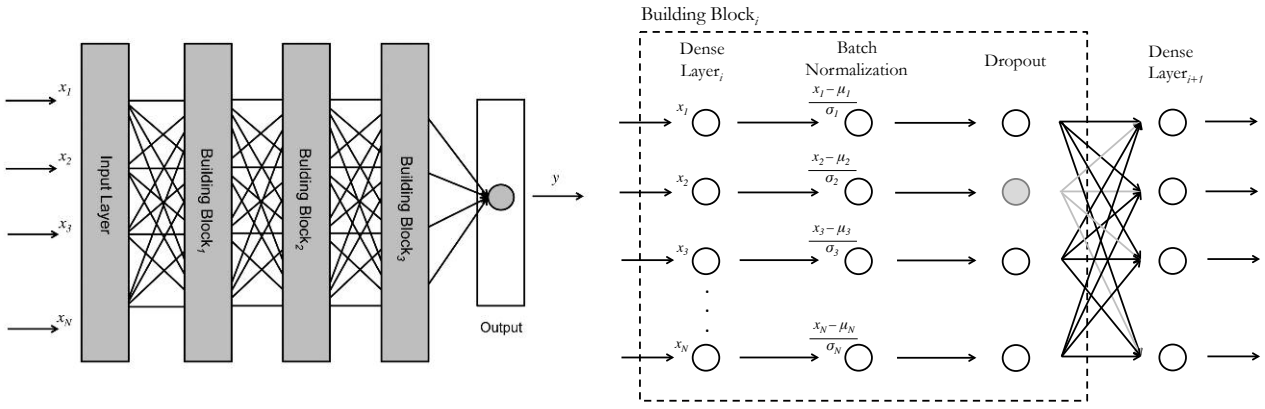


Figure 1: left (a), DNN Architecture; right (b), building block of the DNN model.

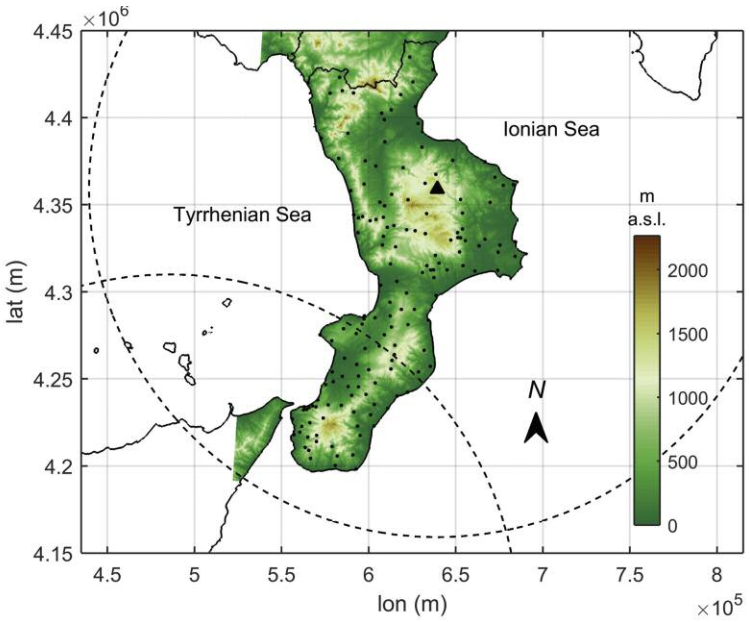


Figure 2: Map of Calabria. The small circles represent the rain gauges, and the large dashed circles the radar ranges.

Rain gauge data are provided by the Calabrian DCP, weather radar data are delivered by the Italian DCP and MSG data are acquired and decoded by a DVB receiver station in high rate information transmission (HRIT).

The data analytics module is devoted to training the rainfall estimation model (REM) from the raw data stored in the KB. It includes three sub-modules designed to handle the whole knowledge discovery flow: data preprocessing, data sampling and model building. Some transformation and filtering operations have to be performed before data could be provided as input to the learning algorithm. The data preprocessing module performs the necessary data cleaning methods for handling the different data issues: missing values, outliers, and noisy data.

A classification model based on DNN architectures is used to estimate the rainfall. Specifically, a feed-forward fully-connected neural network including dropout and batch normalisation layers, shown in Figure 1 (a), is employed to provide more accurate predictions for heavy rainfall events. The building block (Figure 1 (b)) composing our architecture includes three base components: (i) a fully-connected dense layer using a rectified linear unit (ReLU) activation function for each node composing the layer, (ii) a batch-normalisation layer for improving stability and performance of the current dense layer, and (iii) a dropout layer to reduce the risk of overfitting. Moreover, we devised a suitable weighted loss to tackle the class unbalancing problem due to the rarity of (dangerous) heavy rainfall events.

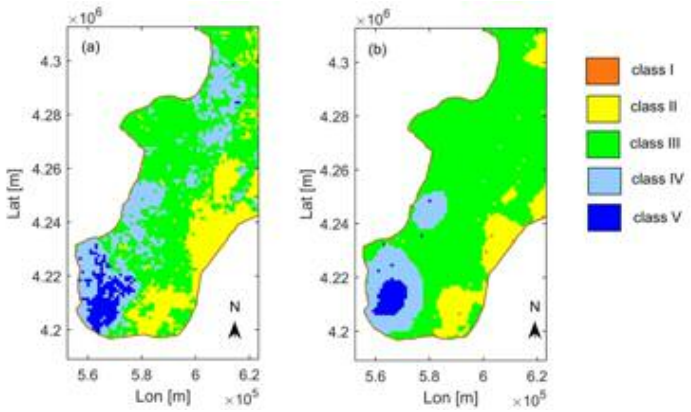


Figure 3: An example of the areal rainfall field estimation for Calabria (left, using our method, right using KED).

To evaluate the solution, we applied it to a real challenging scenario: data from Calabria, a peninsular region in Italy (see Figure 2), provided by the Italian Department of Civil Protection. It represents an effective test case for a number of reasons: despite having more than 700 km of coastline, it is one of the most mountainous regions in Italy, therefore exhibiting strong climatic variability. Furthermore, the interaction between the complex orography, the heat flux from the Mediterranean Sea and the high reliefs near the warm sea, support the convective instability of the region. In addition, floods and landslides are quite frequent here.

Experimental results show significant improvements in comparison with KED (see Figure 3) and with other machine-learning techniques. Although it generate more false alarms, our method detects more rainfall events, in particular for the latter two classes, representing exceptional and/or extreme rainfall events.

Our method is being implemented within the RAMSES (RAilway Meteorological SEcurity System) system, a pilot CNR project, recently co-funded by RFI SpA, that aims to mitigate geo-hydrological risk along the railway [L1].

Link:

[L1] http://www.irpi.cnr.it/project/ramses/

References:

[1] S Gabriele, F. Chiaravalloti, A. Procopio: "Radar–rain-gauge rainfall estimation for hydrological applications in small catchments", Advances in Geosciences, 44, 61-66, 2017.

[2] N. Nanding, M. A. Rico-Ramirez, D. Han: "Comparison of different radar-raingauge rainfall merging techniques", Journal of Hydroinformatics, 17(3), 422-445, 2015.

[3] G. Folino, et al.: "A Deep Learning based architecture for rainfall estimation integrating heterogeneous data sources", IJCNN 2019, pp 1-8, 2019.

Please contact:

Gianluigi Folino
ICAR-CNR, Italy
gianluigi.folino@icar.cnr.it

Can 5G and Machine Learning Replace the Global Positioning System?

by João Gante, Gabriel Falcão (University of Coimbra) and Leonel Sousa (INESC-ID)

Despite being available to civilians since the 1980s, the Global Positioning System is still the standard method for positioning. While unquestionably precise enough for most uses, it requires a dedicated antenna and a significant amount of energy from mobile devices. Using 5G's millimetre wave networks and machine learning, our work shows that we can obtain similar accuracies without these drawbacks.

The advent of 5G is expected to bring new wireless communication capabilities, but it will be accompanied by other challenges. One of 5G's highlights is the introduction of millimetre wave (mmWave) communications, defined by the use of wavelengths between one and ten millimetres, unlocking a significant block of untapped bandwidth. However, with mmWave transmissions, the propagation properties change dramatically: the resulting radiation not only has severe path loss properties, but also reflects on most visible obstacles.

In wireless communication systems, the physical size of the antennae is proportional to the selected wavelength. Therefore, for mmWave systems, whose wavelength is much smaller than most system's, mobile devices can opt between spending less volume or having more antennas. If they follow the latter option, they will have access to beamforming (BF), a signal processing technique that can counteract the aforementioned mmWave drawbacks through a steerable and directive radiation pattern. In fact, when line-of-sight (LOS) communications are unattainable, the focused beam can be aimed towards obstacles, such that its reflection reaches the desired target. As a consequence, multiple signals can co-exist in the same frequency band with reduced interference levels, with careful beamforming execution. No wonder mmWaves are so desirable, at least in theory – they belong to an underutilised part of the spectrum where high levels of spatial multiplexing are possible.

The recent focus in mmWave communications also led to the proposal of new positioning systems. The accuracy achievable in certain conditions is remarkable, having sub-metre precision in indoor and ultra-dense LOS outdoor scenarios. Nevertheless, to be broadly applicable to outdoor localisation, a positioning system must also be able to accurately locate with devices in non-line-of-sight (NLOS) locations, using a limited number of base stations (BS).

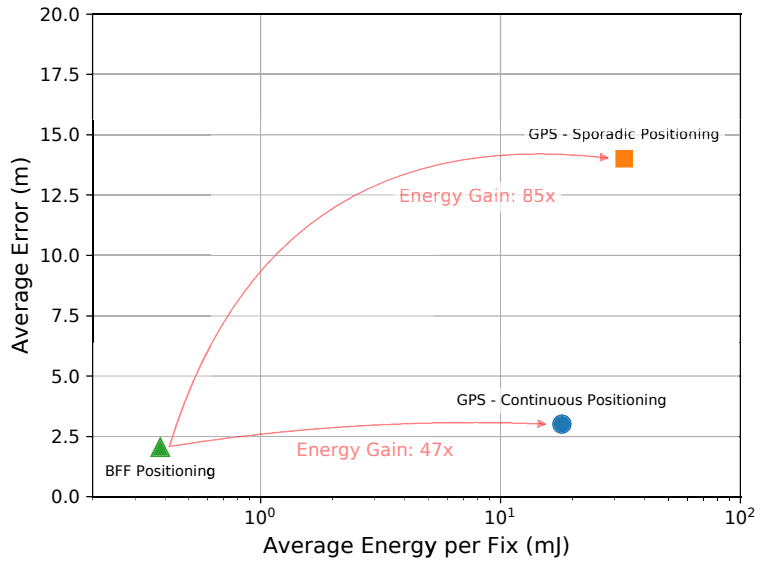


Figure 1: Average error vs average energy required per position fix for the positioning technologies discussed in this article. The proposed BFF positioning system has an accuracy comparable to low-power GPS implementations, while achieving energy efficiency gains exceeding 47× per position fix.

These requirements, allied to multiple, often overlapping non-linear propagation phenomena such as reflections and diffractions, pose serious challenges to the traditional geometry-based positioning methods. In fact, recent mmWave experimental work conducted at New York University demonstrates that geometry-based methods cannot be directly applied to accurately locate NLOS targets, and thus new solutions are needed.

Our team at Instituto Superior Técnico (Lisbon) and Universidade de Coimbra (Coimbra) started addressing this problem in 2018 with the following question: if the BF process can deliver such spatial selectivity, can we use that selectivity to gather spatial information? In mmWave transmissions obstacle interactions are deterministic, resulting in some attenuation and in a change of direction, which impacts the distance travelled by the signal. Given that the time between transmission and reception depends on the distance traversed by the signal, and that 5G is designed for relatively short ranges, it becomes very difficult to sample the signal so as to identify individual interactions, resulting in a significant amount of lost spatial information. However, when BF is available, we can focus the signal in a particular direction, and isolate the interactions from that spatially selected transmission. Repeating this process so as to cover all possible transmission directions, the receiver is able to sample a set of information-rich signals, which we've termed "beamformed fingerprint" (BFF) [1]. With the availability of fingerprint data, machine learning methods and hierarchy techniques were proposed to infer accurate position estimates. Using a single BS, our method achieved Global Positioning System (GPS)-level results for single-point estimates (3.3 m), in a scenario containing mostly NLOS positions, providing posi-

tioning capabilities whenever there is mmWave coverage.

The goal of a positioning system is to estimate the position of a target, which is a direct consequence of its movement. The movement of a user, in turn, is limited by physical restrictions, such as velocity and acceleration, as well as human-made constraints, such as traffic rules. As a consequence, it is possible to leverage additional sources of information if sequences of positions are considered, as opposed to single-point estimates. In [2], we employed temporal convolutional networks (TCNs) when sequences of BFF are available to the system, effectively enabling the system to track a mobile device. Our work with TCNs [L1] achieved the state of the art for NLOS mmWave outdoor positioning, having an average estimation error as low as 1.78 m, with a root mean squared error one order of magnitude smaller than the second-best work for the same problem and, more impressively, more accurate than low-power GPS implementations.

If a positioning method is to displace the GPS as the default positioning method, it must boast similar accuracy levels and lighter hardware and energy requirements. In fact, being a 1980s technology that requires coordination with satellites, the GPS receivers are locked to specific frequencies, which require dedicated antennae, and have power-hungry signal processing requirements. Our most recent work [3] was aimed at answering these practical questions, regarding the BFF positioning system. In essence being a mmWave positioning system, it shares its hardware requirements with mmWave communications, having no additional requirements in mmWave-enabled devices. Finally, the proposed BFF positioning system is also 47 times and 85 times more energy efficient per position fix (for continuous and sporadic fixes, respectively) than low-power GPS implementations, as shown in Figure 1.

Having shown that 5G and machine learning methods can reach GPS-level accuracy levels with lower hardware requirements and much higher energy efficiency, our goal is to advance the industry towards a new paradigm – one that empowers smaller positioning-enabled devices and does not result in space debris.

Links:
[L1]: https://kwz.me/h4x

References:

[1] J. Gante, G. Falcao and L. Sousa: "Beamformed Fingerprint Learning for Accurate Millimeter Wave Positioning", in IEEE VTC2018-Fall.
[2] J. Gante, G. Falcao and L. Sousa: "Deep Learning Architectures for Accurate Millimeter Wave Positioning in 5G", in Neural Processing Letters.
[3] J. Gante, L. Sousa and G. Falcao: "Dethroning GPS: Low-Power Accurate 5G Positioning Systems using Machine Learning", in IEEE JETCAS.

Please contact:

Gabriel Falcão
University of Coimbra, Portugal
gff@co.it.pt

Leonel Sousa
INESC-ID and Universidade de Lisboa, Portugal
las@inesc-id.pt

Faster Flow Predictions with Intrusive Neural Networks

by Yous van Halder and Benjamin Sanderse (CWI)

Numerically solving the Navier-Stokes equations is an important tool in a wide range of industry applications involving fluid flow, but accurately solving them is computationally expensive. Many of these numerical solutions need to be computed, and a trade-off between computational time and accuracy needs to be made. At CWI we developed a method that attains a speed-up of up to 100 times when solving these equations, based on intrusive neural networks.

Recently developed machine-learning based fluid solvers can accurately simulate fluid flows significantly faster than conventional numerical approaches [L1, L2], but they are "non-intrusive", i.e., the flow solver is used as a black-box [1,2]. This non-intrusiveness may result in unsatisfactory extrapolation of the machine learning algorithms outside the training set.

Our intrusive neural network approach uses a machine learning algorithm, a neural network, that is trained using both low- and high-fidelity data and is used intrusively in the solver which stimulates information transfer between the neural network and fluid solver to boost extrapolation capabilities. To be precise, we employ a convolutional/deconvolutional neural network inside an existing numerical fluid solver, which is trained on a set of low- and corresponding high-fidelity solutions in a pre-processing stage. Convolutional layers in the neural network are used to extract non-observable latent quantities, while deconvolutional layers are used to effectively increase the dimensions of the input of the neural network. In our case the considered solver is a Particle In Cell/FLuid Implicit Particle (PIC/FLIP) [L3], which is a combination of a grid-based method, such as finite element, and particle-based method, such as smoothed particle hydrodynamics. The solver evolves a set of particles (representing the fluid) over time by computing fluid velocities on a Cartesian grid, where the fidelity of the simulation is determined by the resolution of the grid and the total number of particles. When a large number of particles is used, the accuracy is determined by the accuracy of the grid veloc-

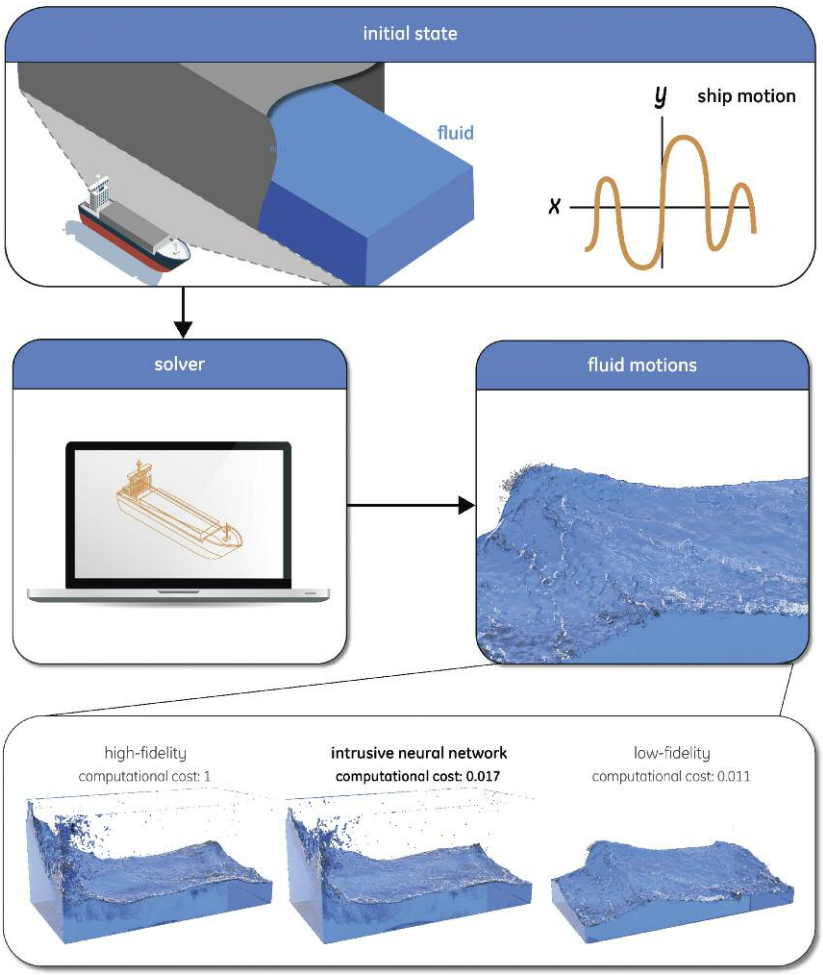


Figure 1: Intrusive neural networks can be used, for instance, to predict sloshing in a very effective and efficient way, making it possible to use it as a predictive tool for decision making. Our new method to solve Navier-Stokes equations for this and other applications in engineering is about 100 times faster than before, whilst maintaining almost the same accuracy. Picture: CWI.

ities, of which the accuracy is determined by the resolution of the grid. A parallel implementation of the PIC/FLIP solver can simulate millions of particles on a coarse grid in real-time and we therefore assume that the fidelity/computational cost is determined by the resolution of the grid. The idea is to run the PIC/FLIP fluid solver on a coarse grid, but to use a neural network inside the solver to enhance the effective resolution of the computational grid to accurately advect the particles. At first sight this might seem an impossible task, but enhancing coarse-grid solutions by using low-level features is in fact not new and is inspired by Large-Eddy Simulation (LES) for turbulence simulations, where the small turbulent scales are modelled using the quantities on the coarse grid.

We demonstrate how the multi-fidelity neural network approach works for simulating sloshing fluids. Sloshing fluids occur for instance when transporting a liquid that is contained in a ship carrier (see Figure 1). The ship's motion induces a sloshing motion of the liquid in the transport tank, resulting in waves that impact the tank hull and may cause failure of the containment structure. A large set of numerical fluid simulations are required to assess the possible danger of liquid sloshing, given the enormous range of possible ship motions. The intrusive neural network approach enables each simulation to be performed at the cost of a low-fidelity simulation, while still obtaining accurate results.

In this case, the deconvolutional neural network is trained on randomly generated low- and high-fidelity fluid sloshing data, which is obtained by creating an ensemble of simulations with random ship motions. This training data then consists of pairs of low- and high-fidelity solutions for which the neural network will serve as the mapping between the two fidelities. After training we can use our approach to enhance a low-fidelity sloshing simulation where the ship motion was different from the motions that were used during training. This gives a clear indication that our approach is indeed able to make predictions outside the training set. Example results are shown in Figure 1.

We clearly see a significant increase in accuracy with respect to the low-fidelity results, while the computational cost of our approach is approximately 100 times smaller than the high-fidelity computational cost. Our approach shows promising results when increasing the accuracy of a wide range of sloshing simulations. However, our approach is not yet able to enhance solutions that involve phenomena that were not encountered in the training set, e.g., obstacles in the flow. We expect that a carefully constructed training set may alleviate this issue.

Links:

[L1] https://kwz.me/h4D
[L2] https://kwz.me/h4F
[L3] https://kwz.me/h4G

References:

[1] L. Ladický, et al.: "Data-driven fluid simulations using regression forests," ACM Trans. Graph., vol. 34, no. 6, pp. 199:1–199:9, 2015.
[2] L. Ladický, et al.: "Physicsforests: real-time fluid simulation using machine learning," in ACM SIGGRAPH 2017, pp. 22–22, 2017.

Please contact:

Yous van Halder
CWI, the Netherlands
y.van.halder@cwi.nl

Benjamin Sanderse
CWI, the Netherlands
b.sanderse@cwi.nl

Surrogating and Calibrating Finite Element Models of Tall Timber Buildings

by Blaž Kurent, Boštjan Brank (University of Ljubljana) and Aleksandar Pavic (University of Exeter)

As the number of tall wooden buildings increases, a good understanding of their dynamic behaviour becomes important. This calls for the collection of empirical data, namely in-situ measured dynamic responses, to enable the calibration of finite element models, the use of surrogates, Bayesian structural identification and uncertainty quantification.

The accuracy of the finite element (FE) models of modern tall timber buildings (TTBs) under service loading (e.g., wind) is poorly understood because of the lack of information and knowledge about stiffness and damping in TTBs. These structural properties are uncertain mainly because of the use of modern timber structural systems and various types of connections between the timber structural elements. Moreover, the wind-induced dynamic excitation is becoming the governing design criteria (besides the earthquake excitation in seismic areas) for determining size and shape of modern TTBs.

The world's tallest timber building, Mjøstårnet in Norway, is 85.4 m high. For timber buildings of this height (and even lower) wind can generate vibrations that cause discomfort or annoyance to occupants due to the perceived horizontal swaying. Wind-induced vibrations can also generate undesirable peak acceleration levels. More experimental data on the dynamics of various types of TTBs under service loadings is required. The DynaTTB research project [L1], funded by the ForestValue research program [L2], performs on-site ambiental vibration tests and forced vibration tests (excited by shakers) on TTBs across Europe. The obtained data provides correlated modal properties (natural frequencies, mode shapes and damping ratios) as references for the design of future TTBs.

These experimental data are not only used to validate the best-engineering-judgement FE models for TTBs but also to improve them. During the calibration of the FE model, chosen (material) parameters of the model are tuned so that the computed response better represents the reality of TTBs' dynamics. If the modelling error and the discretization error of the initial FE model are small, the results of the FE model calibration give insight into the more realistic values of the chosen parameters. To gain further insight into the parameter values, the Bayesian probabilistic framework is used. One of the methods is stochastic Monte Carlo sampling, which needs tens or hundreds of thousands of evaluations of the FE model. To reduce the computational time, a surrogate of the FE model may be built using machine learning tools. The simplest method for constructing a surrogate is the response surface method, where simple polynomial functions are fitted to the FE model results at certain design points. A more advanced method is generalised polynomial chaos expansion, where probability distributions of input parameters determine a set of orthogonal polynomials, evaluation points and their weights for fitting. Other advanced methods include kriging, neural networks, Bayesian networks and other techniques [1].

Let us briefly present the results of the calibration of the FE model of a seven-storey timber building located in Glasgow, UK, that is made entirely of cross-laminated timber (CLT) panels. Applying forced vibration modal testing [2], eight vibration modes were identified. The initial best-engineering-judgement FE model (with CLT described by the layered and orthotropic shell finite elements) reasonably matched the first four experimental modes, which is an acceptable model for predicting the structural dynamics in low frequencies. However, to get a better FE model that covers higher frequency response, a model calibration using a genetic algorithm was applied. After performing calibration of the initial finite element model with six material parameters being tuned, we obtained two additional matching modes and improved the accuracy of all computed natural frequencies signifi-

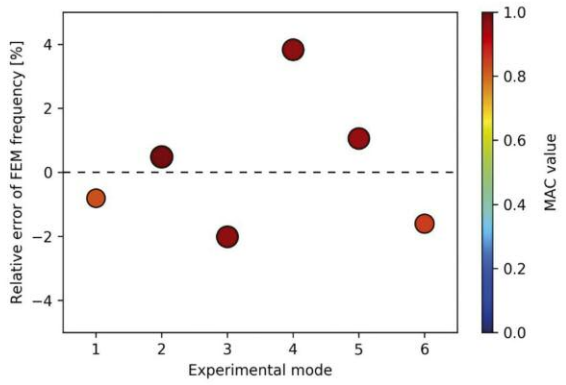


Figure 1: Results obtained by the calibrated finite element model in comparison with the experimental data: relative error in frequency and modal assurance criterion (MAC) value.

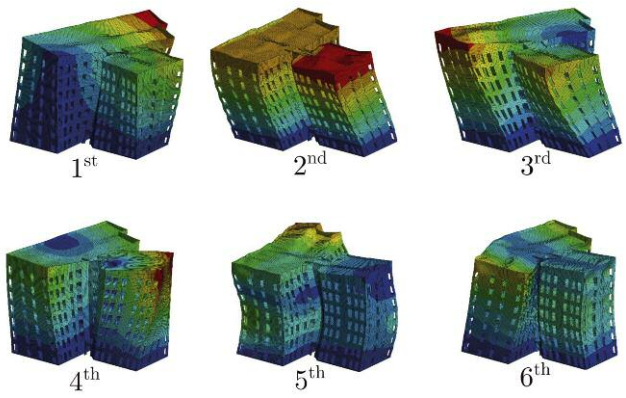


Figure 2: First six mode shape of the seven-storey timber building in Glasgow, UK, made entirely of cross-laminated timber panels.

cantly. Comparison of the computed natural frequencies and vibration modes with the experimental results is shown in Figure 1, and the first six mode shapes are presented in Figure 2. The abovementioned calibration process needed 485 evaluations of the FE model before it converged. We also tested the FE model calibration by using the surrogate, which was built with kriging on 89 evaluations of the finite element model, saving more than 80% of the computational time. For the presented example, the results of calibrations were comparable, but there is a concern about the reliability of surrogate models for more nonlinear systems that require more extensive model fitting and cross-validation. For this building, a Bayesian FE model calibration using a polynomial chaos expansion based surrogate model is in progress. The uncertainty quantification will also be performed.

In conclusion, the finite element model calibration, the surrogate modelling, the Bayesian structural identification and the uncertainty quantification are being applied to some of Europe's tallest timber buildings within the framework of the DynaTTB research programme [3], in order to get a better understanding of their dynamic behaviour and provide reference for the design of future TTBs.

Links:

[L1] https://www.dynattb.com/

[L2] https://forestvalue.org/

References:

[1] H. G. Matthies: "Uncertainty quantification with stochastic finite elements", Encyclopedia of computational mechanics, 2004.

[2] W. K. Ao, A. Pavic: "FRF-based modal testing of sway modes using OCXO synchronised accelerometers for simultaneous force and response measurements", EURODYN 2020 conference proceedings (in press).

[3] R. Abrahamsen et al.: "Dynamic response of tall timber buildings under service load – the DynaTTB research program", EURODYN 2020 conference proceedings (in press).

Please contact:

Boštjan Brank
University of Ljubljana, Slovenia
bbrank@fgg.uni-lj.si

Aleksandar Pavic, University of Exeter, UK
A.Pavic@exeter.ac.uk

Low-Dimensional Flow Models from High-Dimensional Flow Data with Machine Learning and First Principles

by Nan Deng (IMSIA, ENSTA Paris, IP Paris & LIMSI, UPSaclay), Luc R. Pastur (IMSIA, ENSTA Paris, IP Paris) and Bernd R. Noack (Harbin Institute of Technology)

Reduced-order modelling and system identification can help us figure out the elementary degrees of freedom and the underlying mechanisms from the high-dimensional and nonlinear dynamics of fluid flow. Machine learning has brought new opportunities to these two processes and is revolutionising traditional methods. We show a framework to obtain a sparse human-interpretable model from complex high-dimensional data using machine learning and first principles.

The complexity of fluid flow dynamics comes from high-dimensional and nonlinear dynamics. However, in many cases, the complex dynamics can be described by some elementary structures, such as vortical structures in the wake and impinging flows. These typical structures, featuring typical spatial and temporal scales, reveal the underlying mechanisms hidden by high dimensionality. Benefiting from the powerful feature extraction ability of machine learning, dimensionality reduction has become much easier, and numerous methods have been developed. Meanwhile, nonlinear system identification has become more flexible and intelligent. The dynamics can be derived not only from traditional approaches (stability analysis, Galerkin projection, etc.), but also from pure data-driven regression methods. A combination of data-driven system identification with the constraints from the traditional method leads to a physics-based reduced-order model (ROM), which provides us with a better understanding of complex flow dynamics and contributes to the design of effective control and optimisation.

We propose a semi-supervised modelling methodology, combining an unsupervised pattern recognition using proper orthogonal decomposition (POD) with a supervised system identification under physics-based constraints. This will obtain a least-order mean-field model [1], which can explain the hidden relation between the fluctuation and the mean flow fields. Based on the classic Galerkin framework, a critical optimisation relies on considering the symmetry of the mean flow field and the anti-symmetry of the fluctuation flow field, which dramatically simplifies the difficulty of system identification and improves their interpretability.

For instance, this framework is applied to a transient flow of the toy system "fluidic pinball" [L1], exhibiting two successive Hopf and pitchfork bifurcations, as shown in the 3D phase portrait of Figure 1. At Re=80, investigated with direct numerical simulation (DNS), starting close to the symmetric steady solution, the flow state first reaches an unstable limit cycle, associated with a statistically symmetric vortex shedding, then gradually loses the statistical symmetry and eventually reaches one of the two stable limit cycles, associated with a statistically asymmetric vortex shedding.

The reduced-order modelling strategy is illustrated in the lower part of Figure 1. The first step is to determine the least-dimensional manifold considering all the transient dynamics. Following a Galerkin approach, the necessary number of degrees of freedom to describe the Hopf bifurcation is three, while it is two for the pitchfork bifurcation. From mean-field considerations, the instability is triggered by anti-sym-

metric eigenmodes with respect to the x-axis in both the Hopf and the pitchfork bifurcations. By contrast, the distortion of the base flow from the steady solution to the mean-field, represented by the "the shift mode", is symmetric and slaved to the corresponding anti-symmetric active modes.

The model decomposition has a clear purpose of flow feature extraction. With snapshots of the velocity field obtained from the DNS, this process can be purely data driven. The anti-symmetric active modes for the Hopf and pitchfork bifurcation are extracted from the permanent asymptotic regime by the two leading POD modes and the difference between the two mirror-conjugated asymmetric steady solutions, respectively. The two slaved shift modes are recognised from the difference between the symmetric steady solution and their mean flow fields. As a combination of the two bifurcations, the Galerkin expansion consists of at least five independent, orthogonal modes, whose dynamics are coupled together when the Reynolds number is far beyond the critical value of the pitchfork bifurcation.

Nonlinear system identification for a five-dimensional system is very challenging. Twenty-five linear terms (l_{ij}) and 75 quadratic terms (q_{ijk}) need to be identified. Based on symmetry considerations, more than half of the terms vanish. The key terms (growth rates, frequency, slaving relations, and nonlinearity parameters) for each bifurcation are identified from the linear stability analysis of the steady solution and from the typical time and amplitude scales of the asymptotic dynamics on the limit cycles. The simplest dynamical system is shown in the bottom-left of Figure 1.

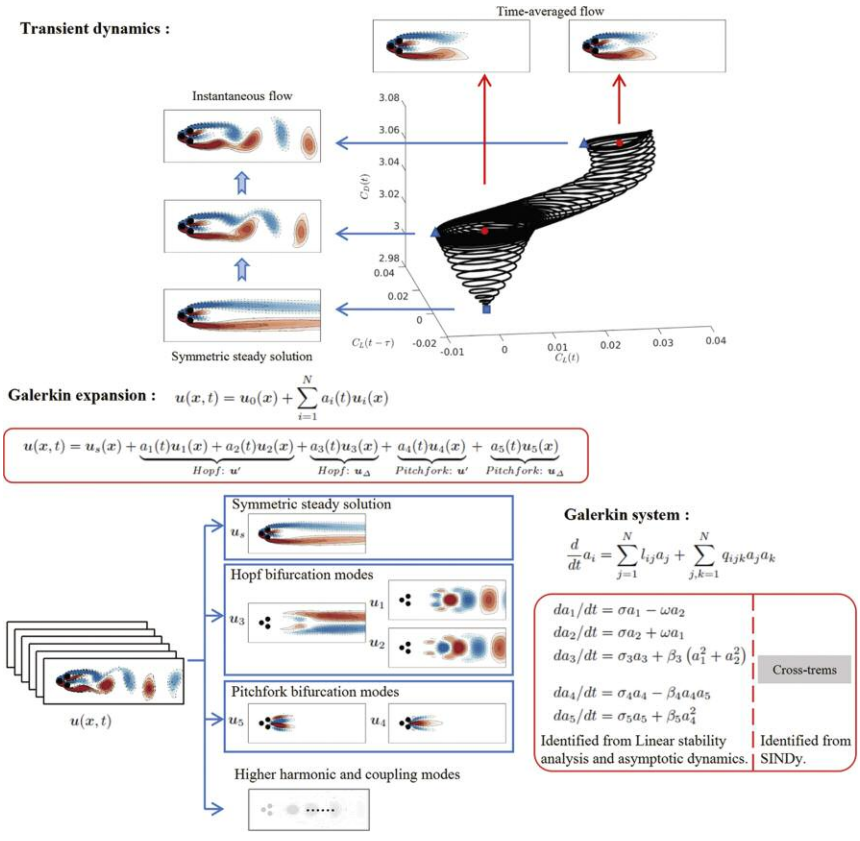


Figure 1: Reduced-order modelling strategy for the fluidic Pinball at Re=80. Upper part: Phase portrait based on the lift coefficient CL(t), lift coefficient with delay CL(t-τ), and drag coefficient CD(t). A transient scenario starting close to the symmetric steady solution (blue square) first reaches a symmetry-centred limit cycle and then loses the statistical symmetry and approaches to an asymmetry-centred limit cycle. For these two limit cycles, the mean flow field is marked with a red point, and a sampled instantaneous flow field is marked with a blue triangle. Lower part: Mean-field Galerkin expansion and corresponding Galerkin system. In the red boxes is the resulting least order model. The modal decomposition is carried out for the two consecutive bifurcations. Both physics-based and data-driven methods are used during the system identification: the key coefficients in the simplest system are identified from the linear stability analysis and the asymptotic dynamics, and the cross-terms are identified from a sparse regression algorithm (SINDy).

The remaining cross-terms are identified with a supervised method, using a sparse regression algorithm under constraints (SINDy) [2].

A sparse, easily interpretable, five-dimensional Galerkin model has been derived from an infinite flow system. The main features of the manifold on which the dynamics take place are correctly identified by this least-order mean-field model. For the interested reader, further details can be found in Deng et al. [1]. Comparing to numerous kinds of ROMs [3], the mean-field model emphasises the nonlinear dynamics of the base-flow distortion from the fluctuation, which provides a theoretical basis from the Reynolds equation for the ROM. The linear-quadratic dynamics of the corresponding Galerkin system is fully consistent with the quadratic nonlinearities of the Navier-Stokes equations.

This framework can be generalised to other flow configurations or even more complex dynamical regimes, like the quasi-periodic regime of the fluidic pinball at higher Reynolds numbers. Based on the least-dimensional ROM, some additional degrees of freedom could be included, e.g., higher harmonic modes, to allow the energy to flow to smaller scales in the model. It can also be used for nonlinear model-based control by including additional actuating modes. Thanks to the clustering and classification abilities of machine learning, the procedure can be further automatised with automated learning of the state space.

Link:

[L1] http://berndnoack.com/

References:

[1] N. Deng, et al.: "Low-order model for successive bifurcations of the fluidic pinball", J. Fluid Mech., 2020.

[2] S. L. Brunton, J. L. Proctor, J. N. Kutz: "Discovering governing equations from data by sparse identification of nonlinear dynamical systems", Proc. Natl. Acad. Sci., 2016.

[3] K. Taira, et al.: "Modal analysis of fluid flows: An overview", AIAA J., 2017.

Please contact:

Nan Deng, IMSIA, ENSTA Paris, IP Paris & LIMSI, UPSaclay, France
nan.deng@ensta-paris.fr

Taming Non-Linear Dynamics and Turbulence with Machine Learning Control

by Guy Y. Cornejo Maceda, François Lusseyran (LIMSI, CNRS, Université Paris-Saclay) and Bernd R. Noack (Harbin Institute of Technology)

Machine learning control is a model-free method based on artificial intelligence techniques to build optimal control laws exploiting non-linear dynamics in an unsupervised way. It is a game changer in discovering new dynamics for experiments and real-life applications.

The idea of controlling fluid flows is not a new one, but it is currently gaining momentum. The intense activity in this area is a consequence of both the technical developments of sensors and actuators and the advances in automatic learning methods. These methods hold potential to address many engineering problems in fields such as transport, industry and energy production. By manipulating the flow, performance can be improved and energy consumption reduced.

There are three main categories of flow manipulation: aerodynamic shape optimisation, passive control and active control. While aerodynamic shaping aims to optimise the shape of the system, passive control adds features to further increase performance, e.g., turbulators on wings and spoilers on sport cars, which are often accompanied by drawbacks such as drag increase. Active control on the other hand, uses actuators, for example Coanda blowers, small jets, plasma actuators and rotating blades, to introduce energy to the flow. Such control requires a positive balance between the energy injected and the energy recuperated but sensors make it possible for these techniques to be adaptable to the flow conditions; this is referred to as closed-loop control.

By adding feedback, closed-loop control tackles inherent challenges for control such as robustness to various conditions, rejection of noise due to the environment and model uncertainty reduction. Closed-loop control can offer enormous benefits: lift increase for drag reduction, gust mitigation, mixing enhancement and noise reduction, to name a few. However, predicting the proper control command is a difficult problem, relying on knowledge of the flow that is usually inaccessible in real time due to the time-delayed response, frequency crosstalk and the inherent high-dimensionality of the flow. Moreover, only partial information about the dynamics is accessible with non-intrusive sensors. Finally, there is currently no precise way to determine the number, location and kind of sensors and actuators required in a closed-loop system; this decision is guided by engineering wisdom.

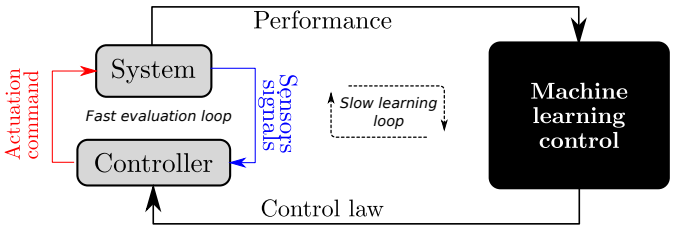


Figure 1: Machine learning control learning process. A fast evaluation loop is used to test the candidate's control laws and a slow learning loop builds new control laws from previous ones, based on their performance.

Linear control theory may have some answers but it calls upon restrictive linear hypotheses that discard non-linear interaction between modes, which is often essential for fluid control. Indeed, examples of wake stabilisation with high-frequency forcing and low-frequency forcing show that frequency crosstalk is a key enabler for control. Thus, it is essential to consider a model-free approach that considers dynamics in its entirety, both linear and non-linear aspects, such as in turbulence [1]. Building an adequate control that makes use of non-linear interactions between modes, exploits time-delays and takes advantage of the physical properties of turbulence is not an easy task and relies on recent advances in machine learning and artificial intelligence. The main approach to solving this near-inextricable control problem is to formulate it as a function optimisation problem where the function to optimise is the control law itself, according to a cost function to minimise. The control law is a mathematical expression function of flow sensors and time-dependent functions such as periodic functions, that commands the actuators. The cost function is a measure of the performance of the control law regarding the objective, e.g., drag power, lift fluctuations, which are defined so as to determine a minimum for the optimal solution. The resulting optimisation problem is another non-convex optimisation problem, which is difficult, but much more accessible to machine learning/artificial intelligence techniques.

Machine learning control (MLC, [2]) is precisely a machine learning solver that solves such hard non-convex optimisation problems and thus builds a controller in a model-free approach. It builds on the pioneering work of Dracopoulous [3] on genetic programming to build a map between the actuation command and the sensors. The learning process mimics the Darwinian principle of survival of the fittest to build fitter control laws by testing successive control laws in an unsupervised way.

Indeed a fast evaluation loop is repeated several times to test candidate solutions, and MLC combines them to produce new ones (slow learning loop), based on the performance of past control laws (Figure 1). The strength of MLC relies on its ability to build non-linear control laws reproducing known control methods including model-based (ERA/OKID), open-loop strategies (multi-frequency forcing), closed-loop strategies (phasor control, ARMAX) and also linear and non-linear combinations of them. It can virtu-

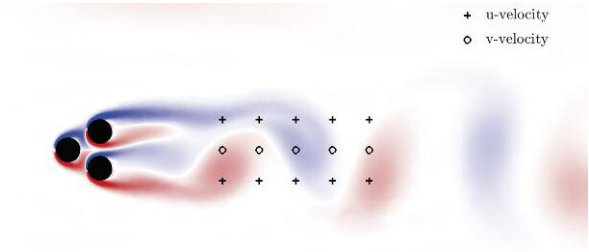


Figure 2: Vorticity field for the unforced natural flow at Reynolds number 100. A grid of 15 sensors downstream with four delays (t, t-T/4, t-/2, t-3T/4), for a total of 60 sensor signals, have been chosen to reduce the net drag power with MLC, with t representing the natural period of vortex shedding. Simulations have been carried out with a DNS solver provided by Marek Morzynski of Poznan University.

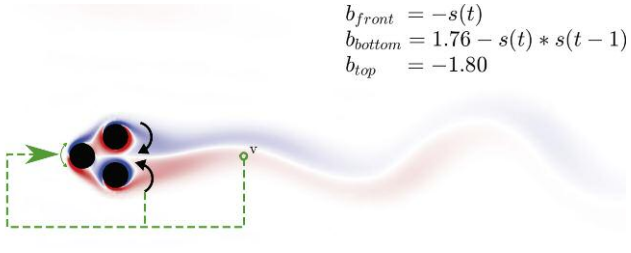


Figure 3: Vorticity field for the best control law found with MLC; a combination of boat-tailing (inward symmetric rotation of the two back cylinders) and phasor control for the front cylinder. This solution reduces the net drag power by 46% and uses only one of the 15 sensors, in this sense, MLC acts as a sensor optimiser.

ally build any function/control law in the control law space.

The aim of this project is to accelerate the learning process of MLC through a smart exploration of the control law space. For this purpose, our MLC code [L1] is benchmarked on a cluster of three equidistant cylinders immersed in an incoming flow called the fluidic pinball [L2]. Thanks to the independent rotation of the cylinders, the fluidic pinball can reproduce up to six wake suppression mechanisms comprising frequency crosstalk mechanisms. MLC achieved a 46.0% net drag reduction of the fluidic pinball by successfully combining two strategies from literature: boat-tailing and phasor control without imposing this prior knowledge in the automated optimization. The unforced flow is shown in Figure 2 and the MLC solution in Figure 3. Current advances show an improvement of the learning rate by a factor of three and with future work, we aim to achieve a factor 10 improvement. Such an acceleration will benefit a multitude of experiments by enabling multi-parameter testing for experiments and simulations with the potential to reveal hidden control mechanisms unreachable with model-based approaches. This approach is a change in the classical paradigm where control laws are derived from the analysis of the system. The machine learning paradigm instead starts from the optimal solution and understanding the mechanisms at play comes afterwards.

Links:

[L1] https://www.cornejomaceda.com/

[L2] http://berndnoack.com/

References:

[1] Y. Zhou et al.: "Artificial intelligence control of a turbulent jet", J. Fluid Mech. 2020 in print.

[2] T. Duriez , S.L. Brunton, B. R. Noack: "Machine Learning Control Taming Nonlinear Dynamics and Turbulence", Springer, 2017.

[3] D. Dracopoulos: "Evolutionary Learning Algorithms for Neural Adaptive Control", Springer 1997.

Please contact:

Bernd R. Noack,Harbin Institute of Technology (Shenzhen), China
bernd.noack@hit.edu.cn

Towards Self-Learnable Software Architectures

by Henry Muccini (University of L'Aquila) and Karthik Vaidhyanathan (Gran Sasso Science Institute)

Software systems are developed following standard architecting practices but are prone to uncertainties that result in suboptimal behaviour in certain unexpected conditions. We humans learn by making mistakes, by adapting to environments, situations, and conditions. What if our software architectures could automatically learn to handle uncertainties? Just like self-driving cars, self-learnable software architectures have the potential to outperform current software in unanticipated circumstances.

It is currently an exciting time for software sciences, with the rapid advances that are taking place in computing. Software has already impacted the lives of billions of people across the world, and with developments in AI, future software is set to solve more complex challenges. However, the more complex challenges they solve, the more challenging it is to architect and maintain these systems. The heterogeneous composition of modern software systems contributes to this complexity. Moreover, these systems are subjected to various uncertainties at run-time such as application downtime due to high CPU utilisation, server outages, etc. These can have a big impact on the quality of service (QoS) offered by the system, thereby impacting the experience of the end-user.

Modern systems mitigate these issues by using AI techniques at the application level or by using "self-adaptation", which allows these systems to recover quickly in the event of failures/downtime. However, these issues often give rise to costly maintenance works in the system architecture. This calls for better mechanisms that can foresee any possible performance issues and avoid manual maintenance by intelligently modifying the architecture itself at runtime. To this end, our research goal is to create self-learnable software architectures that can foresee possible issues, autonomously perform the required patchwork and learn from the experience to intelligently improve the architecture over time. This shall be achieved using a combination of deep learning

and reinforcement learning techniques based on the run-time QoS data. The approach will enable software architects to avoid costly maintenance cycles, thereby allowing them to create reliable and self-sustaining software systems.

Our approach can be integrated with any running software system. Figure 1 shows a high-level process pipeline from our approach.

Once the architecture has been validated and deployed, at execution time the monitor process keeps a check on the QoS metrics of the system such as utilisation, throughput and response time. This process is accomplished by keeping track of the system execution logs. These real-time data are then stored in a database such as Elasticsearch through a process of gathering QoS metrics with the help of technologies such as Apache Kafka. The gathered data are further processed to extract actionable insights from the learning process. The learning process uses two types of machine learning techniques. It first uses deep neural networks to forecast the expected QoS metrics of the system. Based on this forecast, it uses model-free reinforcement learning techniques like Q-Learning to select the best decision to reconfigure/adapt the architecture. This decision is further communicated to the "adapt" process, which adapts the architecture. The adaptation involves performing actions such as adding/reducing execution memory, replacing faulty components, adding/removing instances in the case of a microservice architecture, reconfiguring various execution parameters, etc. In order to build neural network models to perform forecasts, the learning process keeps training the neural network at regular intervals with the QoS metric data collected. This is done to avoid errors in prediction.

Since no machine learning process is 100% accurate, the adapted architecture undergoes a "validate" process that checks if the modified architecture guarantees the QoS requirements. Furthermore, the deploy process deploys the architecture. After every deployment, the variation in QoS metrics is used to measure the quality of adaptation. This is used as feedback to the machine learning process for further improvement. This process continues throughout the software lifecycle. In this manner, the approach ensures that any possible QoS issues are identified at a much earlier stage and the software architecture is modified so as to prevent any system halts thus ensuring reliability and thereby avoiding any costly maintenance cycles. Moreover, over time, the approach ensures that the architecture automatically learns to better handle uncertainties.

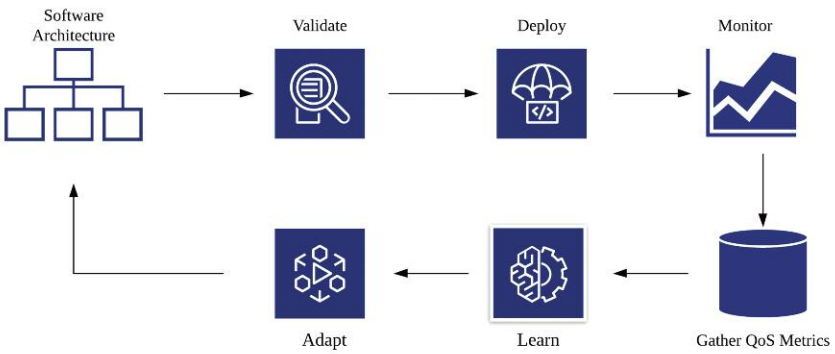


Figure 1: Overall process pipeline of a self-learnable architecture. This process keeps running throughout the software lifecycle to ensure that the architecture is able to continuously learn and improve by itself to handle the different possible uncertainties that might affect the QoS of the system.

As a first step, we have developed an approach that enables a given IoT architecture to automatically learn and improve its QoS, in particular, energy consumption and data traffic throughout its lifecycle. Further, our experimentation on the simulated version of the IoT system under development for European Researchers Night at L'Aquila, Italy [L1] shows that our approach can improve the energy efficiency of a given IoT system by 20% without affecting the system's performance [1]. This has also been developed into a tool, Archlearner [L2] [2] using enterprise-grade big data stack. An integrated verification mechanism using probabilistic model checking techniques ensures correctness of the machine learning technique [3].

We are currently extending our approach to traditional microservice-based systems as well as to systems based on serverless computing. This is a step towards our bigger vision of creating self-learnable software architectures. The initial results are promising, giving us confidence that we will be able to make this into working reality.

Links:

[L1] https://nottedeiricercatoriaq.it/

[L2] https://mysat.gitlab.io/archlearner-web/

References:

[1] H. Muccini, K. Vaidhyanathan: "Leveraging Machine Learning Techniques for Architecting Self-Adaptive IoT Systems", Proc. of the 6th IEEE Int. Conf. on Smart Computing (SMARTCOMP 2020), to appear.

[2] H. Muccini, K. Vaidhyanathan: "ArchLearner: leveraging machine-learning techniques for proactive architectural adaptation", in Proc. of the 13th European Conf. on Software Architecture (ECSA), 2019.

[3] J. Cámara Moreno, H. Muccini, K. Vaidhyanathan: "Quantitative Verification-Aided Machine Learning: A Mixed-Method Approach for Architecting Self-Adaptive IoT Systems", in 2020 IEEE International Conference on Software Architecture (ICSA), 2020, pp. 11-22, DOI: 10.1109/ICSA47634.2020.00010.

Please contact:

Henry Muccini
University of L'Aquila, Italy
henry.muccini@univaq.it

Karthik Vaidhyanathan
Gran Sasso Science Institute, Italy
karthik.vaidhyanathan@gssi.it

Probabilistic Characterisation of Acoustic and Seismic Signals

by Costas Smaragdakis and Michael I. Taroudakis (University of Crete and IACM-FORTH)

The analysis of long time series of measured acoustic or seismic signals may lead to the extraction/determination of specific features that characterise the signals and the information they carry. Two scientific fields that could make extensive use of signal characterisation are acoustical oceanography, where the acoustic signal can be used as a monitoring tool of the marine environment and seismology in which the seismic signals are rich in information about the geological structure of the earth. We are developing alternative tools for signal characterisation based on a time-frequency analysis of the corresponding recordings followed by a probabilistic feature extraction driven by the hidden Markov theory, a well-known machine learning approach for describing sequential data.

The probabilistic characterisation of signals has been applied, to date, with simulated and real data for problems of acoustical oceanography related to ocean acoustic tomography or geoacoustic inversions [1],[2].

In typical applications of acoustical oceanography a signal of known type is emitted from a sound source and is recorded at some distance from it. The determination of the signal features are used as the first step of an inversion procedure leading to the estimation of the environmental parameters of interest; typically the sound speed structure in the water column and the sea-bed properties, using an appropriate underwater sound propagation model. These parameters are the necessary input data for a variety of models that address the dynamics of the marine environment. The topic of this article is feature extraction/signal characterisation using machine learning techniques.

The signal feature extraction procedure presented here, "probabilistic signal characterisation scheme" (PSCS), consists of two stages. In the first stage, the signal is decomposed into several time-frequency layers using the stationary wavelet packet transform. This analysis provides a well translated-invariant time-frequency representation of the signals, revealing hidden trends and patterns of the data.

In the sequel, the energy significant wavelet coefficients are modelled by a left-to-right Hidden Markov Model (HMM) with Gaussian emission distributions. Such models introduce a hidden variable for each time step and this variable is considered responsible for generating the observed data (wavelet coefficients in our case). Each hidden variable can take one of a predefined number of different values (states). The optimal number of hidden states is calculated using the "Bayesian Information Criterion" (BIC). Thus, an acoustic signal is characterised by means of the parameters determining the HMM. These parameters are the transition matrix and the parameters of the emission distributions. Signal similarity measures are obtained using the Kullback-Leibler divergence (KLD) between the representative HMMs for replica and actual earthquake measurements.

The same procedure is used with seismic signals. Most earthquakes are triggered by a slip over a fault area in the Earth's crust. Probabilistic feature extraction of the seismographs included in some data base of a certain region (representative history of earthquake activity in the region) may lead to the association of the triggered faults with single plane simulated faults. Just like the problems of acoustical oceanography, an appropriate earthquake propagation model should be applied to

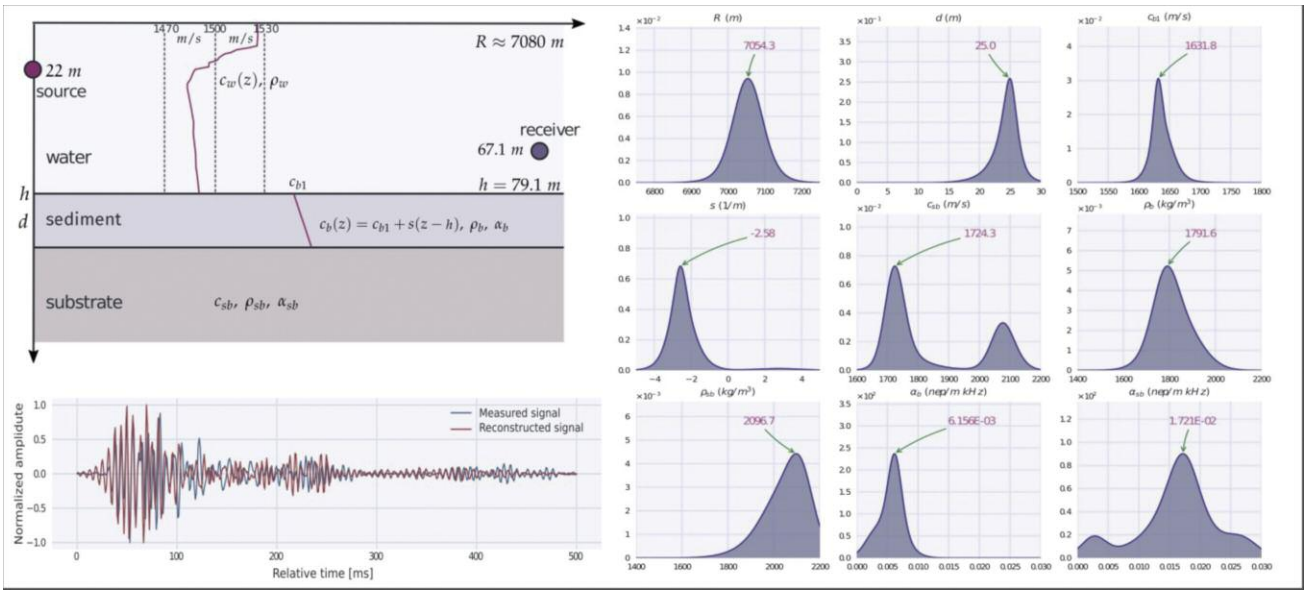


Figure 1: The modelled marine environment of the SW06 experiment, the inversion results in terms of the posterior distributions of the model parameters, and the measured and reconstructed signals.

obtain simulated signals which will also be characterised using the HMM. A similarity measure capable of quantifying the differences between the probabilistic features of each replica signal with the corresponding features of the measured signal of the actual ground motion is once more obtained by the Kullback-Leibler divergence (KLD).

The set of the optimal simulated source mechanism of actual earthquakes can be then used to acquire knowledge about the earthquake cycle (motifs) in a specific region, or cluster of regions.

As an example of the use of the probabilistic signal characterisation scheme in acoustical oceanography, we present a geoacoustic inversion problem of recovering the sea-bed properties of a marine environment using a single recording of the acoustic field due to a known underwater sound source. Figure 1 shows a model of the marine environment corresponding to the experimental track of the SW06 experiment, an underwater acoustics experiment held in the Atlantic Ocean, close to the New Jersey shore. The experiment involved light bulbs as acoustic sources imploded at a depth of 22 m and recorded at a depth of approximately 78 m and range of 7 km. Figure 1 presents the inversion results obtained after applying the PSCS scheme to the recorded and replica signals and using a genetic algorithm over a wide search space on the sea-bed parameters shown in the figure as symbols (not values). These are: The sound speed and density of the sediment layers, the thickness of the sediment layer and the sound attenuation coefficients, with the range R being an additional unknown. The results are expressed by the posterior distributions of the unknown model parameters over the final population of the genetic algorithm. Finally, Figure 1 provides the comparison of the measured and the reconstructed signals by exploiting the values which maximize the marginal posterior distributions.

Our next research goals include the introduction of a complete toolbox based on a state-of-the-art model-based reinforcement learning (RL) technique for providing solutions to a wide class of problems related to the acoustic monitoring of the marine environment, and the risk assessment of the seismicity in selected areas of Hellenic with high earthquake activity.

References:

[1] C. Smaragdakis: "Acoustic Signal Characterization using Hidden Markov Models with applications in Acoustical Oceanography", PhD Thesis, Heraklion, 2019.

[2] C. Smaragdakis and M. Taroudakis: "Hidden Markov Models feature extraction for inverting underwater acoustic signals using wavelet packet coefficients" in Proceedings Euroregio 2016, Porto, 2016.

Please contact:

Costas Smaragdakis
IACM/FORTH and University of Crete, Greece
kesmarag@uoc.gr

Michael I. Taroudakis
IACM/FORTH and University of Crete, Greece
taroud@iacm.forth.gr

AI Marketplace – The Ecosystem for Artificial Intelligence in Product Creation

by Ruslan Bernijazov (Fraunhofer IEM), Leon Özcan and Roman Dumitrescu (University of Paderborn)

Artificial Intelligence (AI) is one of the key technologies of the future and can provide substantial efficiency and productivity gains for product creation. However, manufacturing companies often lack sufficient expertise to take advantage of AI's potential. The AI marketplace will address this challenge by creating an ecosystem for artificial intelligence in product creation.

Artificial intelligence (AI) is one of the key technologies of the near future. According to a recent policy on AI by the European Commission, AI has become "…an area of strategic importance and a key driver of economic development." [L1]. A recent study by the Bitkom association also revealed that 78% of German companies see AI as a decisive technology for the competitiveness of German industry in the future [L2]. AI offers great potential to improve efficiency and productivity, especially for knowledge-intensive activities, like product creation. For example, AI solutions can automate engineering activities, support engineers in their daily work, and automatically discover new insights from the various data sources in engineering [1, 2]. This can lead to an increase in development capacity, shorter development times and contribute to new innovations in product creation.

Although AI holds great potential for product creation, manufacturing companies are being slow to adopt it. We investigated the impediments for adoption [3] by conducting interviews, workshops and surveys with manufacturing companies as well as with AI solution providers. The main challenges that we identified are summarised in Figure 1.

The main challenge for manufacturing companies is often a lack of AI expertise to help the company identify and utilise the potential of AI. Another important concern is uncertainty around data protection and data security measures employed by AI solution providers, which results in low levels of cooperation between manufacturing companies and AI providers. In turn, AI solution providers lack the required domain knowledge to assess how their solutions might benefit product creation. Moreover, AI solution providers often suffer from insufficient access to engineering data to train and test their solutions as well as from poor access to decision makers in industry.

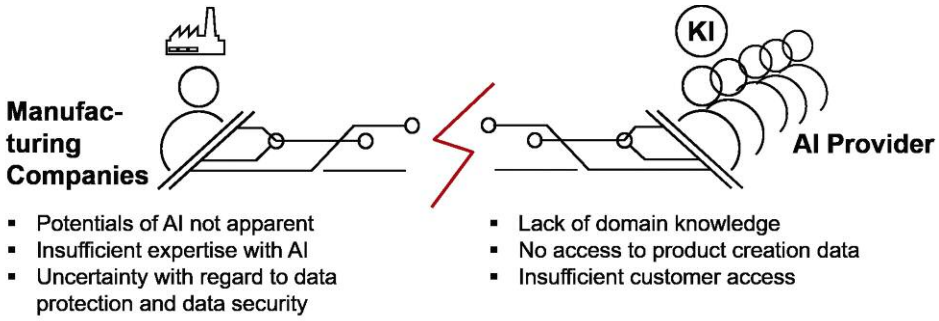


Figure 1: Impediments to the adoption of AI in product creation.

To tackle these challenges, we initiated the project "AI Marketplace" with the goal of creating an ecosystem for artificial intelligence in product creation in the scope of the competition "Artificial intelligence as a driver for economically relevant ecosystems" of the Federal Ministry for Economic Affairs and Energy in 2019. The idea of the project is to build the ecosystem around a digital platform that brings together AI experts, AI developers and manufacturing companies to promote cooperation and joint innovations. The project started in 2020 and has a total runtime of three years. The consortium consists of 18 partners from academia as well as industry and is supported by a strong network of associated partners.

We plan to develop the platform in several stages:

Development stage 1: Intelligent matchmaking

In the first development stage, we focus on matchmaking between manufacturing companies and AI solution providers in order to leverage the realisation of concrete AI use cases for product creation. Manufacturing companies and AI solution providers will be able to describe their main challenges and competencies in their own domain language by means of profile pages on the platform. Based on this information, the platform will suggest project compositions and potential partners. Moreover, the platform will provide a variety of interaction mechanisms, like open challenges and makeathons, in order to support cooperation between manufacturing companies and AI solution providers.

Development stage 2: Protected data space for product creation

The goal of the second development stage is to support the exchange of engineering data between manufacturing companies and AI solution providers by means of a protected data space. This data space will employ existing technologies of the involved partners in order to allow manufacturing companies to share their data in a controlled way. Moreover, we will provide best practices for processing of engineering data to support the development of AI solutions for product creation.

Development stage 3: App store for AI solutions in product creation

The third development stage will extend the functionality of the platform by providing mechanisms for the exchange of ready-to-use AI solutions and services. This will allow solution providers to distribute existing AI solutions directly over the platform to a broad range of potential customers. Moreover, the platform will provide tools and guidelines, like frameworks, libraries, and development methodologies to support the development of AI solutions for product creation.

Development stage 4: Dynamic configuration of AI

The fourth development stage will additionally support a dynamic configuration and integration of existing AI building blocks based on company-specific requirements. This will allow manufacturing companies to combine the solutions of different providers based on their specific needs. For this development stage we will build upon the results of the collaborative research centre (CRC) 901 which has been developing concepts for an on-the-fly configuration of IT-services since 2011 [L3].

Ecosystem services

We are also developing a variety of general services for the ecosystem, like potential analyses and quick-checks to support the onboarding of new members of the ecosystem. More information about the available services can be found on our webpage [L4].

The project AI marketplace is funded by the German Federal Ministry for Economic Affairs and Energy [grant number: 01MK20007A]. The responsibility for the content of this article lies with the authors.

Links:

[L1] https://kwz.me/h4J
[L2] https://kwz.me/h4M
[L3] https://sfb901.uni-paderborn.de/
[L4] https://www.ki-marktplatz.com

References:

[1] L. Bretz, et al.: "Engineering Intelligence - KI-Kompetenz wird für Entwickler immer wichtiger", 2018. https://www.it-production.com/produktentwicklung/ki-kompetenz-entwickler/ (accessed: 12.05.2020).
[2] T. McDermott, et al.: "AI4SE and SE4AI: A Research Roadmap", 2020, in: INSIGHT Magazine.
[3] R. Dumitrescu, M. Drewel, T. Falkowski: "KI-Marktplatz: Das Ökosystem für Künstliche Intelligenz in der Produktentstehung", in: ZWF – Zeitschrift für wirtschaftlichen Fabrikbetrieb.

Please contact:

Leon Özcan
University of Paderborn, Germany
leon.oezcan@hni.upb.de

Reinforcement Learning for Short-Term Production Scheduling with Sequence-Dependent Setup Waste

by Vladimir Samsonov (Cybernetics Lab IMA & RWTH Aachen University), Mohamed Behery and Gerhard Lakemeyer (RWTH Aachen University)

Continually refined and adjusted methods for production planning are among the cornerstones of manufacturing excellence. Heuristics and metaheuristic methods developed to address these tasks are often hard to deploy or lead to suboptimal results under constantly changing conditions combined with short response times of modern production planning. Within the DFG-funded Cluster of Excellence "Internet of Production", a team of researchers from RWTH Aachen University is investigating the use of novel deep learning algorithms to facilitate complex decision-making processes along the manufacturing chain.

Modern manufacturing is a highly complex and dynamic international "ecosystem". Every company involved has to interact with multiple players and fulfil numerous requirements and constraints while compromising between opposing goals. The main task of production planning is to reach and maintain a balance between these constantly changing factors.

Scheduling the orders to production machines in a way that ensures high machine utilisation rates, does not exceed available production capacities, minimises the capital costs, and meets delivery dates turns into a challenging combinatorial task. The size of the solution space grows exponentially with the increasing number of orders to be manufactured and quickly surpasses human capabilities. Time limitations represent an additional dimension of complexity in the case of short-term planning. Multiple events, such as machine breakdowns, change in production priorities, material availability, or personnel shortages can hardly be foreseen and require a quick change of the entire production plan. This calls for approaches that can find solutions to complex combinatorial tasks fulfilling multiple goals and constraints within a short time window available for the decision.

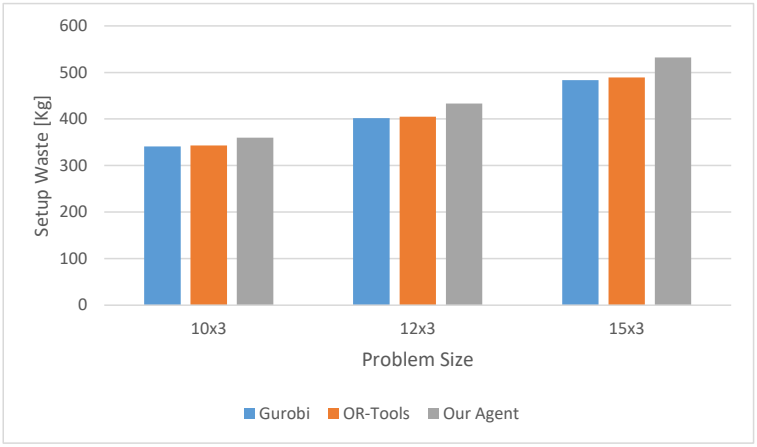


Figure 1: Average Setup Waste [kg], less is better.

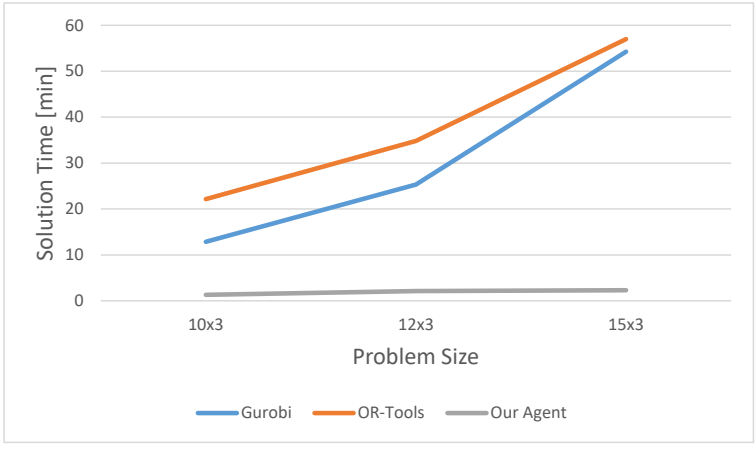


Figure 2: Average Scheduling Time Comparison [mins], less is better.

Deep learning demonstrates a number of successful applications for solving complex tasks with big solution spaces, such as defeating the world champion in the complex game Go [1] or the emergence of the new field "Neural Combinatorial Optimization" addressing combinatorial tasks [2]. We adopt a reinforcement learning approach [3] to the task of weekly production scheduling for foil extrusion. In this case, along with the aforementioned constraints, the setup waste is heavily dependent on the production sequence.

We use historical production data and machine learning to learn complex setup dependencies between thousands of foil types. A trained regression model approximates the setup waste for previously unseen product combinations and serves as a cost function while building a new production schedule. As a part of the validation, dependencies learned by the regression model are extracted with the help of machine learning interpretability methods and are validated through expert knowledge. The trained reinforcement learning agent is benchmarked against two established solvers: Gurobi and Google OR-Tools. These solvers are based on established exact and metaheuristic approaches respectively. Our first validation and comparison runs involve 2,000 different production scenarios of relatively small problem sizes. They show interesting insights into the strength and weaknesses of involved scheduling methods. Figure 1 demonstrates the average resulting setup waste for exemplary planning tasks involving three extrusion machines and ten, twelve, and fifteen

orders referred to as 10x3, 12x3, and 15x3. Figure 2 demonstrates the time each of the approaches requires to find a scheduling solution.

For the considered production scenarios, metaheuristics and exact solvers are able to find solutions that are 1.13% to 10% closer to the optimal solution than our approach. Nevertheless, our approach is able to solve each problem instance at least 20 times faster. For example, on the demonstrated 15x3 scheduling task, exact and metaheuristic methods take 54 and 57 minutes respectively. The reinforcement learning agent solves the same task in under 2.3 minutes. This is a crucial advantage while reacting to an unexpected production deviation.

To conclude, our results encourage the use of reinforcement learning for the task of short-term production planning and scheduling. Future work involves closing the demonstrated optimality gap, extending the use of neural combinatorial optimisation methods to multi-step job shop production environments, as well as working on the aspects of continuous learning, validation, safety, and decision transparency of trained reinforcement learning agents for the deployment in real manufacturing environments.

References:

[1] D. Silver, et al.: "Mastering the game of go without human knowledge", Nature, 2017.

[2] I. Bello, et al.: "Neural combinatorial optimization with reinforcement learning", arXiv preprint arXiv:1611.09940, 2016.

[3] M. Nazari, et al.: "Reinforcement learning for solving the vehicle routing problem", NIPS, 2018.

Please contact:

Vladimir Samsonov
Cybernetics Lab IMA & IfU, RWTH Aachen University
vladimir.samsonov@ima-ifu.rwth-aachen.de

Gerhard Lakemeyer
KBSG, RWTH Aachen University
gerhard@kbsg.rwth-aachen.de

Deep Embedded Vision Using Sparse Convolutional Neural Networks

by Vassilis Pikoulis, Christos Mavrokefalidis (ISI, ATHENA R.C.), Georgios Keramidas (Think Silicon S.A. and Aristotle University of Thessaloniki, Greece), Michael Birbas (University of Patras) and Nikos Tsafas (University of Patras) and Aris S. Lalos (ISI, ATHENA R.C.)

The DEEP-EVIoT project focuses on providing tools to help execute deep multimodal algorithms for scene analysis on embedded heterogeneous platforms (consisting of commercial embedded GPUs as well as dedicated hardware accelerators).

There has been a recent surge in interest in computer vision applications for embedded, portable IoT devices, which satisfy high performance, low computing cost, and small storage requirements. When deep learning (DN) approaches are used, the performance demands and the underlying memory and power requirements of such systems are increased. Furthermore, DN algorithms follow a massively parallel and distributed computation paradigm as they emulate a very large number of neurons operating in parallel.

Deep neural networks (DNNs) have established themselves as prominent tools for solving machine learning (ML) and artificial intelligence (AI) problems (e.g., scene analysis and driver state monitoring), achieving unprecedented results in various applications, and even exceeding the accuracy of human experts in certain classification tasks [1]. However, the quality achieved by DNNs in image/vision applications depends heavily on their size, leading to high computational and storage requirements. This is especially true for DNNs designed to solve demanding image/vision tasks (e.g., 60 million parameters were used in [2])—although these requirements are usually tackled via high-performance computing platforms that include discrete graphical processing units (GPUs). Given the time and computational constraints and the high prediction accuracy required to address potentially life-threatening situations, designing DNNs that can meet the requirements set by the application is not straightforward.

In the DEEP-EVIoT project, we take a twofold approach. First, our target is to design and implement a suite of model compressions and acceleration (MCA) techniques whose goal is to reduce the computational requirements of pre-trained networks, while maintaining their prediction accuracy within acceptable margins. Specifically, we explore techniques based on parameter pruning and sharing, which have been shown to achieve significant MCA with minimal prediction accuracy degradation. Pruning techniques systematically remove unimportant parts of the model (e.g., weights, filters), while sharing techniques reduce the number of operations via subspace clustering and low-rank decomposition, for example, so as to exploit filter redundancies in each DNN layer. Second, we take advantage of the fact that DNNs are very static workloads. Once a DNN has been trained, pruned and quantized, the execution data between the layers remains known and most importantly, is predictable. This characteristic allows hardware and system designers to fine-tune architectures and run-time systems to efficiently execute convolution neural networks (CNNs) (see Figure 1).

Another goal of the project is to design and implement a heterogeneous platform consisting of multicore and multi-threaded embedded, low power GPUs [6] as well as dedicated fixed-logic hardware accelerators with capabilities

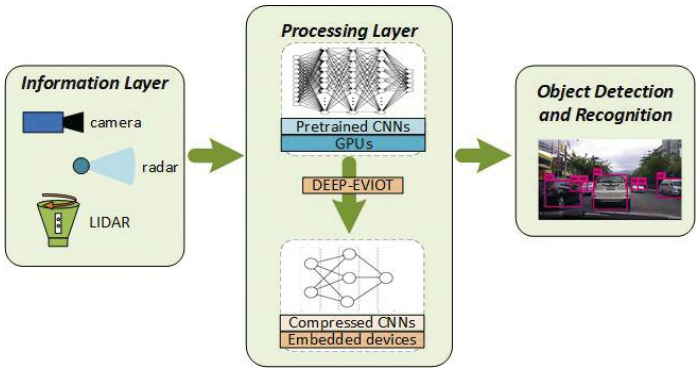


Figure 1: Generation of execution profiles for deep multimodal fusion techniques that trade off performance, energy efficiency, and responsiveness based on hardware availability.

to efficiently execute CNN algorithms. The prime target is the computations performed by the convolution layers. This is because the convolution layers are more complex than both the classification stages and the layers used to down-sample the visual features (e.g., pooling and RELU layers). However, accelerating CNN algorithms is not an easy task, especially when targeting devices with scarce memory and computational resources (e.g., wearable and IoT devices). As part of the project, the majority of the acceleration capabilities will rely on the embedded GPUs offered by Think Silicon, S.A [L2]. Furthermore, a software SDK for vision applications using optimised deep sparse coding techniques, like those mentioned above, especially designed for the platform, will be provided. The SDK will also help the developers to optimally map the CNN layers in a multicore heterogeneous platform consisting of embedded CPUs, GPUs, and dedicated hardware accelerators.

While processors achieve high speeds in sequential algorithms, they have trouble efficiently managing parallel algorithms. In modern CNNs, convolution layers, which are inherently parallelizable, account for more than 90% of the processing workload. They are, therefore, an ideal candidate for the design of specialised hardware [3]. The proposed design, shown in Figure 2, has been built to benefit from the inherently parallel nature of the convolutional layers. Its purpose is to act as a co-processor (an AXI IP Core) which computes the convolutions while the ARM Processor handles the sequential steps of the algorithm. The system reads from and writes to the main system memory using a DMA architecture. The required data is then stored locally in the programmable logic where an array of systolic kernels computes the convolution in parallel thus drastically accelerating the computation of the CNN convolution layers. Due to the systolic nature of the kernels, computations are performed without the need for large or multiple data buses. Once operations are complete, the DMA is responsible for the write-back of the result in system memory. The systolic kernels are designed to take advantage of the data flow (convolutions are done by "sliding" the window over the data) and the reuse of the data (see Figure 3 for an example of a 3 x 3 kernel). Each of the multiply add elements accumulate the new weight and data product with the sum of the previous element. Once the sums of the filtered rows are computed they are added together and these values can be stored in programmable logic memory for use in subsequent steps, and are ready to be loaded into main system memory via the DMA. The system is controlled by a Mealy state machine to account for pipeline delays, row changes and general algorithmic quirks.

The significant benefits of the proposed system extend to a very wide range of applications, from industrial robotics to autonomous vehicles, smart security cameras and the military.

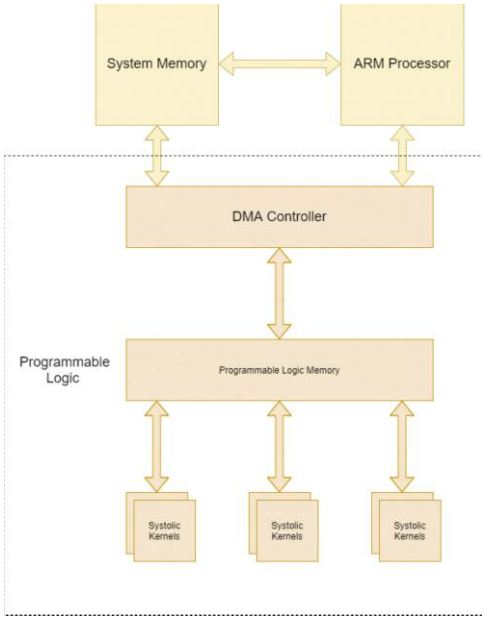


Figure 2: System Level architecture of the proposed solution.

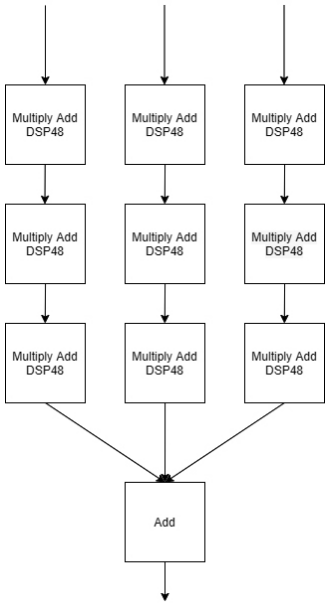


Figure 3: 3x3 Systolic kernel example.

Links:

[L1] https://kwz.me/h4Q

[L2] www.think-silicon.com

References:

[1] V. Sze, et al.: "Efficient Processing of Deep Neural Networks: A Tutorial and Survey", Proceedings of the IEEE, vol. 105, 12, 2017.

[2] A. Krizhevsky, I. Sutskever, G. E. Hinton: "ImageNet Classification with Deep Convolutional Neural Networks," in NIPS, 2012.

[3] Y. Umuroglu, et al.: "FINN: A framework for fast, scalable binarized neural network inference", in Proc. of FPGA, pp. 65–74, 2017.

Please contact:

Aris S. Lalos
ISI, Athena Research Centre, Greece
lalos@isi.gr

Managing Duck Curve Type Energy Imbalances with Variational Recurrent Autoencoder-Based Clustering

by Alkiviadis Savvopoulos, Christos Alexakos and Athanasios Kalogeras (Industrial Systems Institute ATHENA Research Center)

Peak residential energy demand does not always coincide with peak production times. This energy imbalance is known as the "duck curve". Variational recurrent autoencoders can normalise the duck curve, optimise consumption profile clustering, and acquire useful insights for managing energy demand.

The "duck curve" problem is a phenomenon that occurs in energy consumption, with significant imbalances occurring between renewable energy production yield and peak energy demand, owing to their short or non-existent temporal correlation [1]. The term "duck curve" refers to the collective visualisation of the production of renewable energy, the actual energy demand, and the energy supply from non-renewable sources, all as a function of time throughout the day. An abrupt demand surge coincides with the end of daily "green" solar energy production.

Strategies to solve this problem tend to focus on energy storage and management demand. However, algorithmic approaches may accelerate such efforts. The duck curve may be regarded as a specification of a unit commitment problem, where each building connected to the grid contributes to the formulation of the daily period of peak energy demand. Nevertheless, since not all energy grid participants' daily peak demand coincides with total aggregate peak demand, the problem can be mitigated through algorithmic disambiguation between grid users, with demand peaking during the high renewable energy supply period and users heavily tolling the grid during aggregate high demand hours.

The proposed method analyses time series corresponding to residential building energy consumption profiles to determine their load demand characteristics. The sheer volume of data and exceptionally high sampling rate pose a challenge for data handling and analysis. With the aim of achieving an acceptable trade-off between dimensionality reduction and reconstruction loss, we elected to compress the data. Variational autoencoders of the unsupervised learning algorithmic category represent a compression algorithm, utilising state-of-the-art solutions with renown abilities in robust pattern handling of vast amounts of datasets such as neural networks [2]. Moreover, the sequential nature of energy consumption profiles, whose values are partially dependent on the immediately preceding values, means that, due to the aforementioned temporal dependency, steep fluctuations are alleviated mitigated. We address this with long-short term memory (LSTM) neural networks, which take the place of the hidden layers inside the encoding and decoding modules of the variational autoencoders, to handle non-stationary input data and non-coherent energy patterns. The total process of the proposed method is shown in Figure 1.

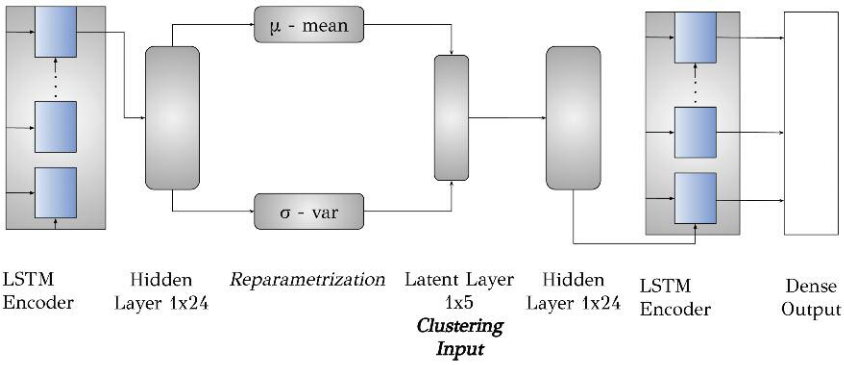


Figure 1: Variational recurrent neural network layer stack.

After the schema hyper-parameters are fine-tuned, the variational autoencoders are deployed, with an outer layer of 48 nodes corresponding to the half-hourly sampling rate of the dataset used, a hidden layer of 24 dimensions, and a double latent layer of five dimensions. This is where the reparameterisation technique of the variational autoencoders takes place, leading to the final compressed representation. The decoding module is symmetric and contains the same layer dimensions. Adaptation of the LSTM inside the schema, mandates the replacement of the classic feed forward neural networks, making the last layer of the LSTM stack, the previous hidden layer that feeds the latent. Similarly, for decoding, the LSTM feeds a final outer dense layer.

After ensuring a much smaller feature space through the encoding step of the algorithmic pipeline, the analysis and clustering of these encoded daily load profiles can be implemented at much greater velocities. The next step of the procedure is to divide the compressed energy consumption profiles into different groups with similar characteristics. The clustering method utilised, in this work which representings this work's its main contribution, is the creation of custom cluster representatives as a reference point for each encoded time series. The fundamental concept of this method is that in order to disambiguate between residential building consumption patterns, two clusters are created: one that correlates with daily patterns whose peak demand resides during the hours of high renewable energy supply, and a second that comprises all time series not compliant with "green" energy.

In order to create these representationsves, which will be compared to each compressed load profile and compute their corresponding similarities, daily solar radiance is taken into consid-

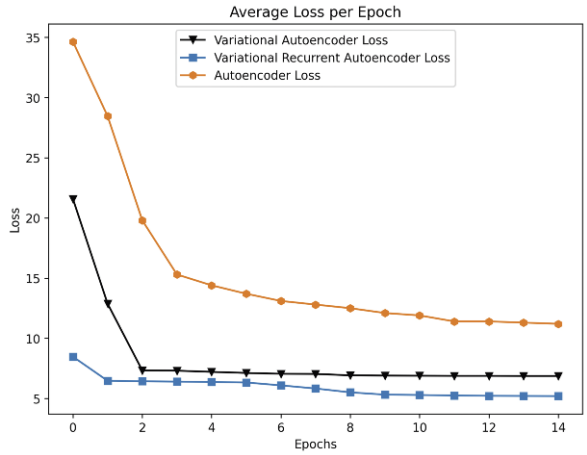


Figure 2: Neural network training results comparison.

eration. The first cluster is constructed maintaining the dataset scaled median value as its maximum, throughout the first morning hours until late afternoon, when solar output is at its highest. The second cluster possesses values at a similar scale, in a sense, complementary, as the evening and night hours must be represented. Then, vectors are encoded, scaled and compared with each time series, so as to determine its class and assign to the corresponding cluster.

Using daily energy consumption measurements of 5,567 London Households [L1] as experimental data, variational recurrent autoencoders outperformed both simple variational autoencoders with classic encoding modules, and vanilla autoencoders without reparametrisation, in terms of loss during training. In order to introduce experimental robustness, the aforementioned neural networks were also tested; however, due to the structure of the household energy data, the proposed schema captured time-dependent intricacies more efficiently, as seen in Figure 2. With reference to clustering results, as depicted in Figure 3, our method performed an efficient separation of time series according to peak demand, into a renewable energy compliant cluster, and a second cluster containing load profiles that must be dealt with in order to solve the duck curve issue.

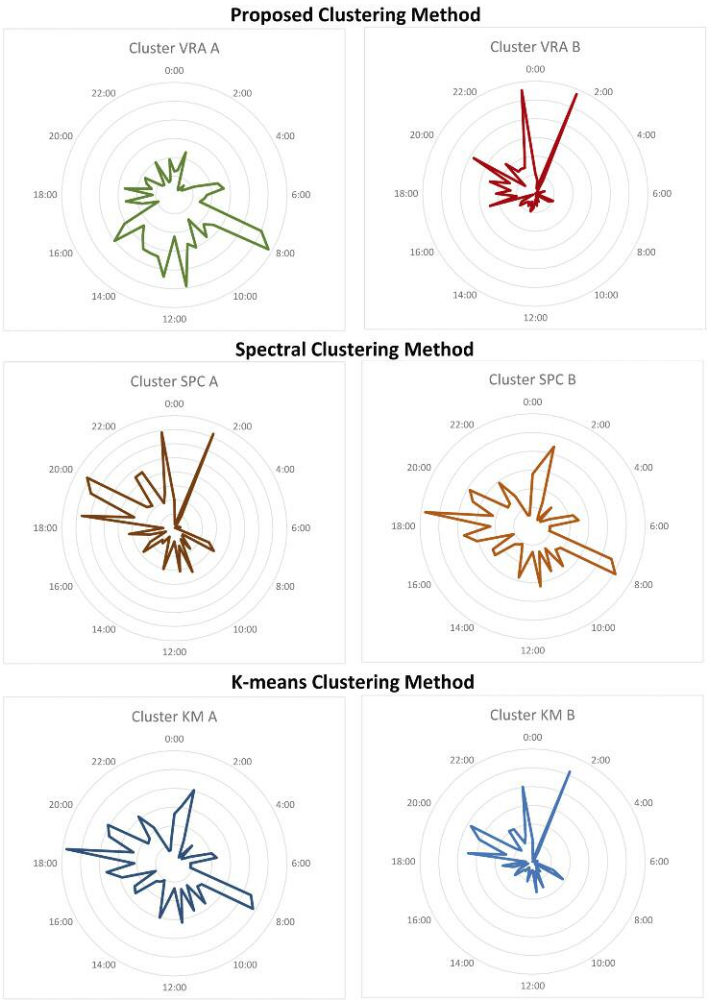


Figure 3: : Experimental clustering results. The figure presents the categorisation of the analysed time series in two clusters, a "green" compatible cluster and a non-compatible cluster, according to their consumption peak times, utilising three approaches: the proposed VRNNL, the spectral and k.-means algorithms.the spectral and k.-means algorithms.

In summary, the proposed algorithm clusters the dataset's input points according to their peak demand time index throughout the day, while simultaneously outperforming classic clustering approaches such K-means and spectral clustering. The algorithm does not surpass traditional clustering techniques in terms of classic metrics such as precision and recall. Rather, its strength lies in its ability to create load profile groups that can produce meaningful insights, and efficiently detect whether a profile correlates with renewable energy production standards, in real time.

We acknowledge support of this work by the project "Enabling Smarter City in the MED area through Networking-ESMARTCITY" (3MED171.1M2022), which is implemented under the "Interreg Mediterranean" programme, co-financed by the European Regional Development Fund (ERDF), Instrument For Pre-Accession Assistance (IPA) and national sources.

Link:

[L1]: https://kwz.me/h4j

References:

[1] H.O.R. Howlader, et al.: "Optimal thermal unit commitment for solving duck curve problem by introducing CSP, PSH and demand response". IEEE Access, 2018.

[2] D. P. Kingma, & M. Welling: "Auto-encoding variational bayes", ICLR2014.

Please contact:

Alkiviadis Savvopoulos
Industrial Systems Institute, ATHENA Research Center, Greece
savvopoulos@isi.gr

Optimization of a Chemical Process with Soft-Sensing Technologies

by Enrique Garcia-Ceja, Åsmund Hugo, Brice Morin (SINTEF) and Per Olav Hansen (Unger)

Process optimisation within industry can reduce production times, as well as material and energy consumption, which translates to more efficient use of resources and money. In the chemical production industry, soft-sensing and machine learning technologies can help to optimise processes.

Digitalisation and automation technologies have permeated industrial processes, and are used, for example, to reduce waste and increase product quality and production rates. These benefits, which have huge impacts on user satisfaction and the environment, are achieved largely through "optimisation"; the efficient use of resources such as time, materials and machines.

In recent years, a new approach, "soft-sensing", has been used to foster optimisation. Soft-sensing uses easy-to-measure variables to predict other variables that are impractical or more expensive to sense (Figure 1). This scenario is common in chemical processes that often take place in harsh environments in which monitoring may be hazardous and require highly specialised and expensive sensors [1]. In this context, soft-sensing technologies can replace expensive monitoring tasks with software algorithms fed from different data sources, for example, sensor readings. Soft-sensing relies on two key ideas: data and prediction. Machine learning is thus an ideal candidate to support soft-sensing operations.

We trialled the use of soft-sensing and machine learning to optimise a chemical process in Unger Fabrikker, a chemical factory in Norway. Unger produces surfactants: compounds used in personal care and laundry products, such as shampoo, toothpaste, soap, detergents and bleach. Unger uses a sulphonation process to produce surfactant variations appropriate for the final product. When transitioning from one product to another, there is a production gap of about 30 minutes, during which waste is produced and an operator needs to take samples and analyse them manually. The output of this analysis is a neutralisation number (NT) that measures the quality of the product.

Predicting the NT value is an expensive task, so to reduce the 30 minute gap, we used soft-sensing and machine learning as the predictive engine to infer the NT value automatically based on other process parameters, such as amount of air injected to the sulphur oven and converter, current amount of sulphur, quantity of organic material and air temperature. Together with Unger Fabrikker, SINTEF and Østfold University, we conducted experiments where we trained different machine learning regression models including a random forest and a neural network to predict the NT value based on the other process parameters. To evaluate the generalisation performance of the models, we trained them using 70% of the data and evaluated the results on the remaining 30%. The data consisted of 14,252 historical measurements. From our experiments, random forest obtained the best results with a mean absolute error of 0.089 compared to 0.115 with a neural network. It is worth mentioning that the data contains some noisy measurements which were manually identified by an expert and removed before training the models. These results show that the quality control task can be reduced from half an hour to a couple of seconds. The machine learning models have been incorporated into the monitoring system and their results are stored in a database for further validation.

We are currently testing machine learning methods to automatically identify outlier measurements. Preliminary promising results show that it is possible to identify outliers using an ensemble learning approach, that is, combining multiple predictive models. We trained three classifiers that predict whether a point is an outlier or not based on the process parameters. The models were one random forest and two Naïve Bayes classifiers. The final result is obtained by taking the majority vote of the three. This approach was able to detect 82.6% of the outliers. The outlier detector is currently being tested in the real system.

Additionally, we are testing a new near infrared (NIR) sensor developed by Prediktor [L2] to infer the product quality based on its readings. The NIR instrument is installed in the production flow in one of the reactors. Machine learning algorithms are fed with data from the spectrum values captured with the NIR instrument to estimate the product quality. This has the advantage of providing additional information, such as indicating whether the process is in a stable phase or not. Data collected during changeover between products is also analysed. In this changeover phase the NIR instrument

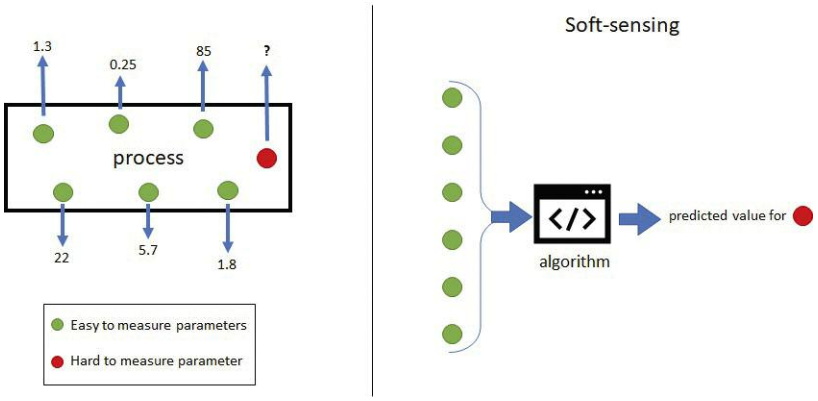


Figure 1: A process with six parameters that are easy to measure and one that is hard to measure (left), and an algorithm that predicts the hard to measure parameter based on the others (right).

provides information that can be used to build a model that could potentially make the changeover between products automatically.

In summary, soft-sensing and machine learning technologies are allowing Unger to perform processes more efficiently and have opened new opportunities to optimise other operations.

The work [2] was conducted as part of one of the use cases of Productive 4.0 [L1], an EU ECSEL project with one of its goals being the design and development of Internet of Things (IoT) technologies, including hardware and software for the digital industry.

Links:
[L1] https://productive40.eu
[L2] https://www.prediktorinstruments.com

References:
[1] S. Zhang, et al.: "Online quality prediction for cobalt oxalate synthesis process using least squares support vector regression approach with dual updating", Control Engineering Practice, 2013.

[2] E. Garcia-Ceja, et al.: "Towards the Automation of a Chemical Sulphonation Process with Machine Learning", ICCMA 2019, IEEE.

Please contact:
Enrique Garcia-Ceja
SINTEF, Norway
enrique.garcia-ceja@sintef.no

Per Olav Hansen
Unger, Norway
Per.Olav.Hansen@unger.no

Machine Learning and Chaos Theory in Agriculture

by Sebastian Raubitzek and Thomas Neubauer (Vienna University of Technology)

Machine learning has found its way into agricultural science for analysis and predictions, e.g., of yield or nitrogen status. Results are encouraging, but predictions in agricultural sciences are still tricky because agriculture is a highly complex system, with outcomes depending on a multitude of complex phenomena, such as weather, irrigation and soil properties. We propose future machine learning research in this sector to consider complex systems (chaos theory) and improve machine learning approaches.

In recent years machine learning has found its way into many areas of science, including physics, biology, finance, medicine, and agricultural sciences. Machine learning models are usually employed for analysis and predictions, such as stock price forecasts or population estimates. The more traditional approach of mechanistic models consists of mathematical machines with specific parameters set to perform well on a given task. Machine learning approaches, in contrast, are driven by historical data.

In agriculture, machine and deep learning have the potential to help forecast yields and nitrogen status [1]. For these tasks, it is necessary to use as many available sources of data as possible, e.g., remotely sensed data or historical yield and soil records. Here, as a rule of thumb, the broader the variety of data, the better. The reason for this is that agricultural production depends on many attributes, such as weather, soil properties, topography, irrigation, and fertiliser management. Results of these predictions are encouraging, but a lack of data and the complexity of the systems make predictions in agriculture difficult, i.e., the performance of the algorithms on test data in this sector thus far is insufficient for many applications.

On the other hand, the study of complex systems and chaos theory peaked in the 20th century. Many tools have been developed to measure the complexity of data and to estimate the chaotic behaviour of systems. Usually, those chaotic systems have many degrees of freedom and, therefore, strongly depend on the initial conditions under study. Since agricultural systems depend on many different attributes, we understand agricultural systems as complex systems (See cf. [2] for applications, such as chaotic plant population dynamics, spatio-temporal dynamics for arable land). Complexity can be found on various scales in agricultural systems, from the biological foundation of a single plant to the interactions between different crops, animals, and humans and the dependence on weather and climate interactions.

It is therefore important to consider the complex nature of the systems under study when doing analysis or predictions. Tools such as the Hurst exponent, the fractal dimension, entropy measures, or the spectrum of Lyapunov exponents may prove useful. These tools, which were developed to analyse chaotic or complex systems, are helpful to characterise and analyse agricultural systems. Methods such as the Hurst exponent are referred to as complexity measures, i.e., measuring the fluctuations or long-term memory of a system. Originally the Hurst exponent was

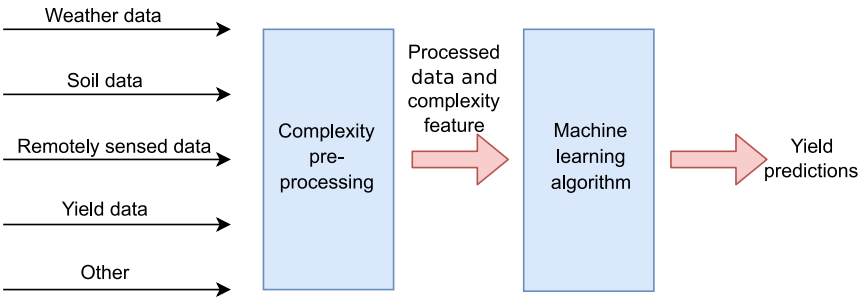


Figure 1: Schematic depiction of the proposed example. A multitude of data is fed into the preprocessing. The complexity-based preprocessing discards noisy data and adds a complexity feature to relevant data. A machine learning algorithm is then employed to predict yields.

invented for use in agriculture, specifically to determine the optimal dam sizes of the river of Nile to guarantee optimal irrigation of the surrounding land. The spectrum of Lyapunov exponents, in contrast, is a measure of the predictability of a system, and it can also estimate the degrees of freedom of a system.

To date, little research has combined the study of complex systems and machine and deep learning approaches, particularly within agricultural sciences. This may be because scientists working in agriculture rarely have expertise in this area of modelling, and likewise many computer scientists, although trained in mathematics, never hear about complex systems. There is usually primary education in physics and mathematics, but a course in complex systems is not part of the curriculum.

In [3], ideas from chaos theory and a neural network were used to predict geomagnetic activity. To be specific, a neural network approach has been improved using the Hölder exponent (Another complexity measure for time series related to the Hurst exponent). There are also some applications for finance where complexity measures have been used to improve predictions or estimate the volatility of data.

As an example of how to combine machine learning and chaos theory, the data stream of the proposed process is depicted in Figure 1. First, all available data sources are pre-processed using tools from chaos theory. In this step, irrelevant or too-noisy data is discarded, and a complexity value for each feature at every data point is calculated. This complexity value at each data point is referred to as a complexity feature. Second, only relevant data and the corresponding complexity features are fed into the machine learning algorithm to calculate yield predictions.

We hope that the examples and references within this article serve to motivate researchers in both computer and agricultural sciences to learn about complex systems, chaos theory, and the corresponding tools. Because of its in-depth treatment of the system under study, a combination of machine learning and chaos theory may increase productivity and sustainability in agriculture.

The authors acknowledge the funding of the project “DiLaAg – Digitalization and Innovation Laboratory in Agricultural Sciences”, by the private foundation "Forum Morgen", the federal state of Lower Austria and by the FFG; Project AI4Cropr, No. 877158.

References:

[1] A. Chlingaryan, S. Sukkarieh, and B. Whelan: “Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review”, Computers and Electronics in Agriculture, 151:61–69, August 2018.

[2] K. Sakai: “Nonlinear Dynamics and Chaos in Agricultural Systems” ,Elsevier Science, Amsterdam, 1 edition, December 2001.

[3] Z. Vörös and D. Jankovičová: “Neural network prediction of geomagnetic activity: a method using local Hölder exponents”, Nonlinear Processes in Geophysics, 9(5/6):425–433, 2002.

Please contact:

Sebastian Raubitzek, Thomas Neubauer, Vienna University of Technology, Austria
Sebastian.Raubitzek@tuwien.ac.at, Thomas.Neubauer@tuwien.ac.at

Advanced Data-Driven Manufacturing

by Théophile Gaudin, Oliver Schilter, Federico Zipoli and Teodoro Laino (IBM Research Europe)

In many material manufacturing processes nowadays a large amount of data is created and stored, often without utilizing them to the full potential because of their complexity. Applying state of the art deep learning techniques can be a powerful tool to extract knowledge out of them allowing to get useful insights. In this work we present autoencoder-based machine learning models to find links among composition, properties and processes applied to two prototypical industrial applications.

In any modern industrial plant, a myriad of data is generated by sensors or by the recording of various input parameters. This stream of data is important for various reasons, including reproducibility and quality control. With abundant and diverse data, machine learning can be used to extract information about the process itself, allowing us not only to predict outcome properties, but also to improve manufacturing parameters. In most cases, some of the input parameters are independent. For instance, a formulation (alloy or polymeric material) with a specific composition can undergo various processes and one single process can be applied to various formulation compositions. Here, we present variational autoencoders to generate continuous representations of the independent parameters in a reduced space, which is called latent space. We use the latent space to perform an optimisation of the parameters that can be conditioned on part of the input (i.e., given the composition of a formulation, finding the parameters that would yield the highest tensile strength). Our framework is illustrated in Figure 1.

An autoencoder consists of two neural networks: an encoder and a decoder. The encoder converts an input into a fixed sized vector called latent representation. The decoder reconstructs the input given the latent representation. Autoencoders are usually trained to minimise the reconstruction error. Kingma and Welling [1] introduced a variational autoencoder where an additional constraint is added to the encoder so that it produces a more robust representation.

The proposed method uses different autoencoders, one for each group of independent parameters. Going back to the formulation example and its charac-

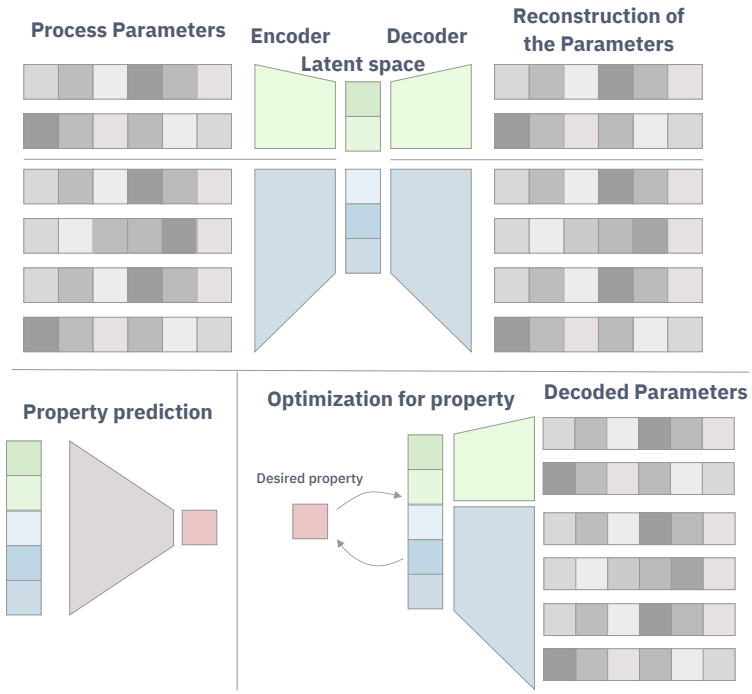


Figure 1: Overview of the model. Process parameters are encoded to the latent space and then reconstructed by a decoder. From the trained latent space, the property can be predicted by an additional neural network. It is also possible to search the latent space to optimise the parameters using a Gaussian process.

teristic processes, we have an autoencoder for the composition of the formulation and another to encode the sequence of processing steps. From these autoencoders we constructed a latent representation from which we can predict various properties (in the case of an alloy, these may be mechanical properties and for polymeric materials, physico-chemical properties) using an extra neural network. For a mere prediction of the formulation properties one could directly use a neural network that takes process parameters as input and outputs the properties values. Instead, inspired by the work of Gomes-Bombarelli et al. [2] in the domain of molecular structure prediction, we opted for the construction of latent spaces, enabling the use of Gaussian processes to identify points that have certain desired properties. These points are then decoded back to the input parameter space. This approach has three advantages. Firstly, it allows mapping a discrete parameter space to a continuous one. Secondly, the latent space has fewer dimensions than the parameter space, making it easier to search. And finally, this approach could lead to solutions that have not yet been observed in the dataset. This flexible framework with different autoencoders also allows the search to be conditioned on a part of the parameters. For instance, we can look for a specific process yielding the highest tensile strength given an alloy composition.

The presented structure is general enough to be adaptable to a variety of material design problems. For example, in addition to modelling the alloying process of metals, the same approach can be used to model an extrusion process for polymers. In this case, our training data consists of formulations and processes, and in some cases also aging conditions. The data was encoded into three latent representations, one for each data type. The trained model for polymers makes it possible to optimise the composition of a recipe for a known process to get a desired property value; analogously, the same model could also optimise a set of processing condition parameters for a given composition. After the optimisation of a property, the model returns points in the latent space representation, which can be decoded either to the corresponding chemical composition of the polymer or to process parameters. These results can be utilised to optimise recipes and fine-tune process parameters.

In conclusion, this data-driven approach to the design and fabrication of novel materials via encoder-decoder based models offers a promising way to compress the data into a reduced latent space to improve the material design task. The benefit of this framework is its general applicability to any kind of data.

References:

[1] D. P. Kingma and M. Welling: "Auto-encoding Variational Bayes", ICRL 2014. https://arxiv.org/abs/1312.6114

[2] R. Gómez-Bombarelli, et al.: "ACS Central Science", 2018, 4, 268–276.

Please contact:

Théophile Gaudin
IBM Research Europe, Switzerland
tga@zurich.ibm.com

Using Deep Learning for Anomaly Detection in Autonomous Systems

by Nikhil Kumar Jha, Sebastian von Enzberg and Michael Hillebrand (Fraunhofer IEM)

The use of deep learning algorithms is largely restricted to application domains where a large amount of labelled data is readily available, e.g., computer vision. Thus, applications of deep learning in autonomous systems for Industry 4.0 are rare. The application of deep learning to anomaly detection within autonomous systems for Industry 4.0 is a current research topic at Fraunhofer IEM. Our latest studies deliver some promising anomaly detection models as well as automated configuration of model hyperparameters.

A system is considered "autonomous" when it makes and executes decisions based on its current state to achieve certain goals, without any human intervention. The research project "KI4AS – Validation of Artificial Immune Systems for Autonomous Systems" focuses on the resilience of autonomous systems by adapting self-healing properties. The use case at hand is a prototype of a mobile robotic system deployed in a smart factory, which travels from one checkpoint to another using autonomous navigation algorithms and heterogeneous sensor data. Being critical in the nature of the operation, mechanisms are needed to ensure system's robustness in unforeseen situations. Data from the onboard sensors can be monitored and analysed in order to detect deviations from normal behaviour. System reconfiguration or recovery methods can be applied subsequently. Anomaly detection methods identify data samples that do not conform with the expected behaviour of the data generating process. These events either implicate immediate serious consequences, or result from a previously unobserved elemental process. Establishing reliable supervised observation rules to detect anomalies is difficult due to the domain complexity and uncertainty. Also, they imply high costs incurred during the modelling of normal behaviour by the domain experts make such model-based solutions obsolete. This work aims at exploring the data-driven methodologies to meet the Industry 4.0 demands of achieving dynamic monitoring of anomalies in real-time autonomous systems while focussing specifically on the deep learning solutions present in the literature.

Recent advances in technology have resulted in an increase in the number of sensors on such autonomous systems, and more frequent sampling. The concomitant increase in the complexity of the underlying patterns in the dataset makes it impossible for conventional machine learning methods to capture the underlying information and structure. Their inability to scale and perform well with the unlabelled, class-imbalanced dataset also poses a serious challenge when it comes to anomaly detection. In this context, we have focused on three areas of research: deep learning methods, feature selection and hyperparameter optimisation.

Deep learning methods are able to robustly learn from large-scale data without needing manual feature engineering. In anomaly detection, they operate with limited label information i.e., knowledge of one-classed "normal" data only. In this case, the deep learning model represents a profile for normal data samples. Patterns that do not conform to the normal behaviour are identified subsequently by explicitly isolating them based on a measure of abnormality (e.g., distance norm). In our current research, we have investigated, evaluated and compared three deep learning models, each offering their own advantages [1]: autoencoder, long

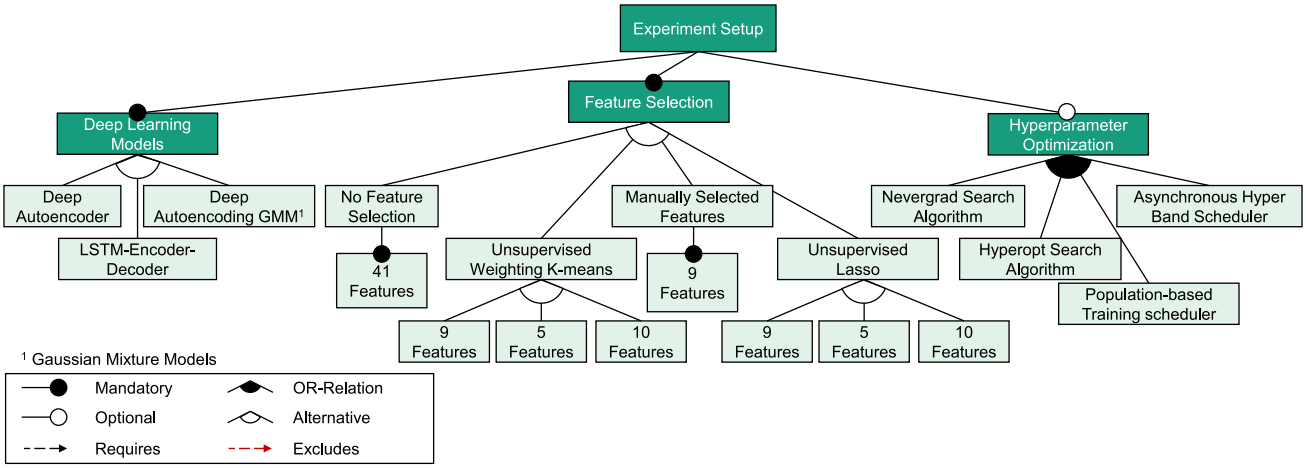


Figure 1: The figure summarises the experiments to evaluate the performance of the models. It drafts the different scheduling and search algorithms used for the hyperparameter optimisation task, and also accounts for the various feature subsets created using the feature selection techniques.

Model	Precision	Recall	F_1-measure	MCC	Inference Time	Memory Usage
Autoencoder	0.62	0.85	0.723	0.727	3.29 secs	79.1 MB
LSTM-ED	**0.565**	**1**	**0.722**	**0.751**	**3.83 secs**	**11.13 MB**
DAGMM	0.324	0.923	0.480	0.545	8 mins 2 secs	75.8 MB

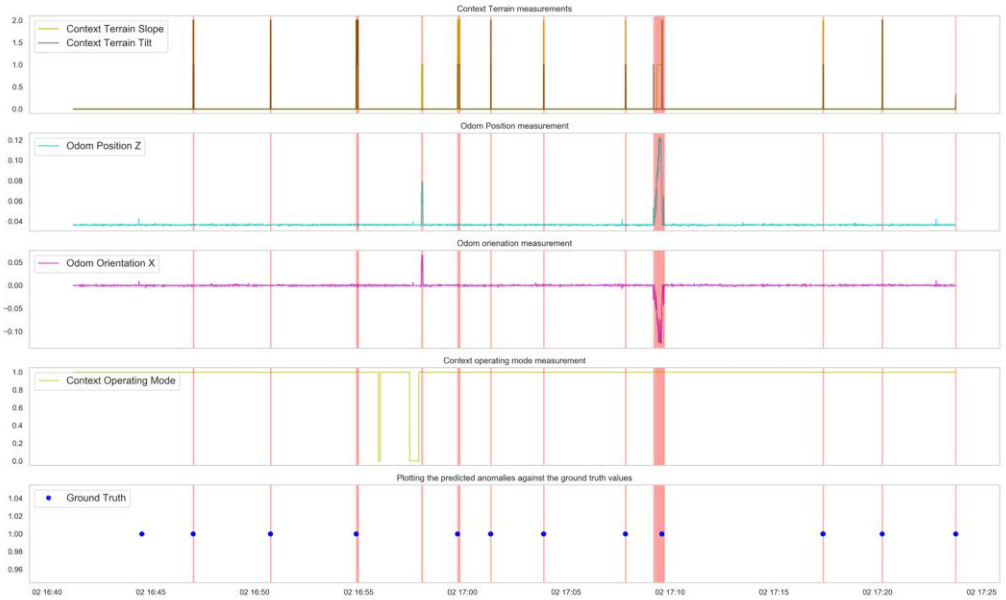


Figure 2: The table shows the quantitative comparison between the most promising configurations for the three deep learning models by using different metrics for classification performance, resource consumption and time taken for inference. Owing to heavy class-imbalance, the metric Matthews correlation coefficient (MCC) is preferred for a reliable statistic. For the best-performing LSTM-ED, the detected anomalies (red vertical lines) are plotted against anomaly ground truth (blue dots) and system signals. Anomalies visibly correlate with the spikes or the irregular behaviour in individual signals.

short-term memory encoder-decoder [L1], and deep autoencoding gaussian mixture model [L2].

When datasets have high dimensionality and sample-sizes, the model learns from, and is thus negatively affected by, extraneous and insignificant features. Feature selection [3] is a way of identifying a subset of important features that can perform the required prediction task, with similar or better performance. For this use-case, we have researched two methods: weighting K-means and LASSO, with weighting K-means showing more promising results.

Another important aspect governing how efficiently the model can generalise over the given data is the selection of an optimal set of hyperparameter setting. Discrete measures in an algorithm like the learning rate, batch size, size and number of layers in the neural network, etc., need finely tuned values to optimally learn the data patterns and optimise the objective function guiding the learning problem. This is achievable by using hyperparameter optimisation [2]. The hand-designed values have always been susceptible to flaws, so automating this process in this work, is the next step in the direction of making the field of machine learning more prosperous and promising.

Several exhaustive experiments on different operating scenarios were made, they are summarised in Figure 1. The primary objective was to evaluate different state-of-the-art algorithms against the anomaly detection task. Figure 2 shows the results of the validated deep learning methods and signal plots for the winning model. The plots help the user to understand and correlate the system behaviour at the time of anomaly detection. The red vertical line represents a detected anomaly, while its width gives an idea about the duration of the detected anomaly. The last subplot is essentially a depiction of the detected and the ground truth anomaly (blue dots).

The LSTM-ED model performed best with an F1-measure of 0.751. Not only did it have the best classification performance, but it is also promising in terms of resource consumption and model inference time. While several state-of-the-art models were evaluated in this work, to our knowledge, no previous study has dealt specifically with the task of anomaly detection using deep learning methodologies on a real-world autonomous system use case. Our work also took the research to a deeper level by combining the unsupervised feature selection techniques along with hyperparameter optimisation performed to determine the best configuration settings and procure competent models.

Links:

[L1] https://arxiv.org/abs/1607.00148

[L2] https://kwz.me/h4I

References:

[1] Chalapathy et al.: "Deep Learning for Anomaly Detection: A Survey" https://arxiv.org/abs/1901.03407

[2] Liaw et al.: "Tune: A Research Platform for Distributed Model Selection and Training" https://arxiv.org/abs/1807.05118

[3] Alelyani et al.: "Feature Selection for Clustering: A Review" http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.8115&rep=rep1&type=pdf

Please contact:

Nikhil Kumar Jha, Sebastian von Enzberg, Michael Hillebrand
Fraunhofer IEM, Germany
nikhil.jha@iem.fraunhofer.de,
sebastian.von.enzberg@iem.fraunhofer.de,
michael.hillebrand@iem.fraunhofer.de

Anomaly Detection on Networks is a Question of Context and Scale

by Leonardo Gutiérrez-Gómez (LIST), Alexandre Bovet and Jean-Charles Delvenne (UCLouvain)

Anomaly detection is an important problem in data mining with diverse applications in multiple domains. Anomalies, also known as outliers, can be defined as individual objects with patterns or behaviours that differ starkly from a background property. Examples of applications include fraud detection in finance, detection of faults in manufacturing, identifying fake news in social media, or web spam detection. Anomalies in real problems may lead to enormous economic, social, or political costs and are often difficult to find, mainly because they are scarce and unknown a priori. Therefore, efficient detection of anomalies may bring significant value to people, companies, and authorities.

In a general form, anomaly detection relies on the ability to characterise and differentiate what is normal from what is unusual or rare. However, such distinction becomes much more challenging when intrinsic interactions between data are present, forming a network of interconnected data.

A network is a mathematical model used to describe complex phenomena in terms of entities and their relationships. Networks are represented mathematically as graphs, with vertices representing entities such as people, products, mobile phones, cities, or neurons; and edges connecting nodes describing pairwise relationships such as friendship (social networks), purchasing (co-purchasing networks), calling (communication networks), physical connection (transportation networks), or anatomical links (brain networks) respectively. Networks are often enriched with attributes or external meta-data on the nodes or edges. For instance, in a social network, we may know demographic information about the people, such as gender, age, nationality, or height. Therefore, meta-data, together with the network structure, provide a more comprehensive description of the problem at hand at the expense of additional complexity.

Finding anomalies on networks with attributes on the nodes is known in the literature as anomaly (or outlier) detection on attributed networks. Some examples include fake profiles in social media networks, review manipulation in e-commerce networks, super-spreaders of infectious deceases in social networks, and brain damage in brain networks. Because the mechanisms of anomaly generation in real-world networks are usually unknown, characterising "ground truth" anomalies is problematic, and therefore estimating and evaluating anomalies is a challenging problem. Thus, anomalous nodes must be defined against a background of "normal" nodes, i.e., the context relevant for an anomaly.

Although many techniques to find anomalous nodes in attributed networks have been proposed, a complete solution has not yet been found. Some approaches define anomalies in local contexts, i.e., community outliers, and others localise anomalies as global outliers, i.e., the context is the entire network. Other authors do not consider any context at all.

However, real networks are characterised by a modular and multi-scalar structure of its components. That is, in real networks, nodes are organised in modules or clusters of strongly connected nodes, with few inter-group connections, e.g., social networks are organised according to people's interests. Besides, networks have multiple levels of abstraction (like Google maps), allowing the network exploration at different scales of detail. A node may be anomalous in a local context but disappears on a global scale, see Figure 1. Thus, it is crucial to take into account the node attributes spanning across the multi-scale structure of networks to correctly define the contexts of anomalies.

To this end, in collaboration with researchers of the Université catholique de Louvain, we provide a more general solution to the problem. We propose a principled way to uncover outlier nodes simultaneously with the context in which they are anomalous, at all relevant scales of the network [1].

Our approach is based on graph signal processing tools [2]. Intuitively, anomalous nodes are those whose attributes differ starkly from the attributes of some context nodes. We characterised anomalies in terms of the concentration of “energy” retained for each node after smoothing specially designed signals localised at each node of the graph. This smoothing operation can be interpreted

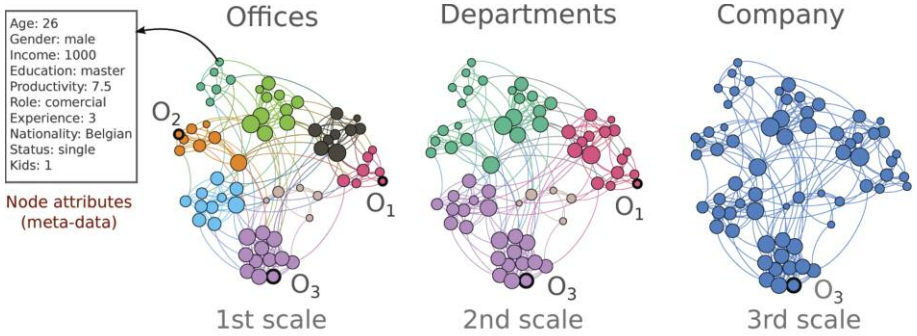


Figure 1: A toy example of a work relation network. Nodes have attributes describing individual features. Node attributes define structural clusters in multiple scales. At the 1st scale, outlier nodes (O1, O2, O3) lie within a local context, i.e., offices. In a 2nd scale, departments emerge as new contexts where O2 is not defined. Finally, at a larger scale, O3 remains as a global anomaly in the context of the whole company.

as a heat diffusion process on the network. That is, heat spreads from one node to its neighbours following the network structure. Also, heat spreads faster between nodes with similar attributes than nodes with dissimilar ones. As a consequence, node attributes of anomalous nodes induce bottlenecks for the heat diffusion. At the beginning of the process, each node has maximum heat because nothing has been spread yet. By increasing the time, i.e, the scale, the heat flows from one node to its neighbours with a propagation rate driven by the similarity between the attributes of the adjacent nodes. The flow tends to remain trapped in regions of the graph where nodes are highly similar or strongly connected. In this way, the diffusion flow unveils the modular composition of the network [3], where for small time scales, fine-grained clusters emerge as local contexts for potential anomalies, and larger times uncover coarser clusters as broader contexts for global outliers. Hence, anomalies are scale-dependent, and the time acts as a zooming parameter (as Google maps), to reveal the context in the network where the anomalies make sense.

Our approach has been validated empirically in synthetic and real-life attributed networks (e-commerce and web spam), outperforming many state-of-the-art methods, with the advantage of being highly efficient and parallelisable. Finally, it is worth noting that our approach is applicable on any attributed network independently of its nature or domain. Moreover, we plan to extend this approach in dynamic time-varying networks, where links between nodes change in time or new nodes join the graph dynamically.

Link:
[L1] https://kwz.me/h0T

References:
[1] L. Gutiérrez-Gómez et al.: "Multi-scale Anomaly Detection on Attributed Networks". Proceedings at 34th AAAI Conference on Artificial Intelligence, 2020.
[2] D. Shuman et al.: "The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains", IEEE Signal Process, 2013.
[3] R. Lambiotte et al.: "Random Walks, Markov Processes, and the Multiscale Modular Organization of Complex Networks", in IEEE Transactions on Network Science and Eng., 2014.

Please contact:
Leonardo Gutiérrez-Gómez
Luxembourg Institute of Science and Technology (LIST), Luxembourg.
leonardo.gutierrez@list.lu

Non-Contact Vital-Sign Monitoring System for Premature Infants in Neonatal Intensive Care Units

by Péter Földesy, Imre Jánoki, Ákos Zarándy (SZTAKI) and Péter Pázmány (Catholic University, Budapest)

A camera and machine learning based system, developed at SZTAKI and by Péter Pázmány at Catholic University Budapest [L1], enables continuous non-contact measurement of respiration and pulse of premature infants. It also performs high precision monitoring, immediate apnoea warnings and logging of motion activity and caring events.

It is essential within hospitals to be able to continuously and reliably monitor vital signs, like the heart rate and respiration of newborn infants—particularly those in neonatal intensive care units (NICU). The heart and respiration rates can be extracted from the electrocardiogram (ECG), which despite being a non-invasive technique still relies on direct contact with the body. The self-adhesive electrodes are relatively expensive, and more importantly they can easily damage the sensitive skin of preterm infants. Therefore, non-contact monitoring is a daily need in NICUs.

Recent studies have shown that non-contact visual vital-sign monitoring is a reliable and accurate technique [1] although, like traditional contact monitoring methods (e.g. ECG, pulse-oximeter), it suffers from motion artefacts. During periods of caring (e.g. baby is removed, skin-to-skin contact with parent visible, cleaning, nurses change feeding tube, etc.) or intense activity the measurements are inaccurate, so these situations need to be treated separately. Our system can detect and handle common activities, such as infant self-motion, phototherapy treatment and low light conditions with infrared illumination, with a high confidence level. Unlike other systems, our system provides continuous monitoring, not limited to motionless periods. The respiration waveform and rate are calculated directly from the chest and abdomen movements, giving a physiologically and computationally more reliable and more established result than extracting them based on remote photoplethysmography (rPPG), like some existing algorithms do.

In the framework of signal and rate extraction, a top classifier runs, with feature extraction and a neural network classifier, distinguishing events and status of the view. This classifier can detect an empty incubator, an active or passive infant, caring and other motion related situations with 98% precision in real life clinical practice. Whenever the infant is detected, the heart and respiration rates are extracted from the video feed as described below.

The heart rate calculation consists of an ensemble of two networks: (i) the signal extractor network, which derives the pulse-signal from the video input; (ii) the rate estimator network, which calculates the heart rate value from the signal. For the former, the PhysNet architecture [2] is applied and the rate estimator is our own network, named RateEstNet. These networks are fused and trained together after the pre-training of PhysNet. We have developed a novel augmentation technique, called frequency augmentation, which produces a uniform heart rate distribution that results in unbiased training (i.e., the network is not biased towards the average heart rate value).

Figure 1: A high-level functional block diagram of our real-time, continuous monitoring system (left), and a typical display view showing the extracted waveforms and rate estimations along with object detection overlay (right).

The respiration rate calculation incorporates the more traditional way of image processing with optical flow and a machine learning approach as well. A dense optical flow algorithm extracts the movements from a series of grayscale images in a time window of about six seconds. The resulting differential sequence is then masked using a U-Net machine learning architecture to filter only the abdomen and chest [3]. The images are summed and processed to get the waveform of the respiration. In a last step, we use a neural network consisting of one-dimensional convolutional layers with a fully connected part at the end to get the respiration rate.

Our system can estimate pulse and respiration rates and can handle medical intervention and heavy motion scenarios built up from an ensemble of hierarchical neural networks (see Figure 1). In physical form the system is under integration into an open incubator pilot product of a leading Hungarian incubator manufacturer, including medically safe night vision illumination and hardware acceleration of the neural networks by a NVIDIA Jetson Nano module.

The method's performance is being evaluated in real-time and on a carefully annotated database collected at the First Department of Neonatology of Paediatrics, Department of Obstetrics and Gynaecology, Semmelweis University, Budapest, Hungary [L2]. The project started in early 2018 and is still running, with further product integration and R&D for extending night vision capabilities, closed incubators, behavioural studies and sleep quality evaluation.

Links:
[L1] https://kwz.me/h4V
[L2] https://kwz.me/h4W

References:
[1] K. Gibson, et al: "Non-contact heart and respiratory rate monitoring of preterm infants based on a computer vision system: a method comparison study", Pediatric research 86.6, pp. 738-741, 2019.
[2] Z. Yu, L. Xiaobai , Z. Guoying: "Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks", Proc. of BMVC, pp. 1-15, 2019.
[3] R. Janssen, W. Wang, A. Moço, G. de Haan: "Video-based respiration monitoring with automatic region of interest detection", Physiological Measurement, vol. 37, no. 1, pp. 100-114, 2015.

Please contact:
Péter Földesy
SZTAKI, Hungary
+36 1 279 6000/7182
foldesy@sztaki.hu

An Automatic Anomaly Detection System (AADS) for Fully Autonomous Ships

by Bekir Sahin and Ahmet Soylu (NTNU)

Various global factors - including, variability in maritime regulations, technological progress, and ecological and environmental problems - have been converging, pointing to the importance of sustainability in the maritime industry. To reduce maritime accidents and the loss of life and property, sustainability needs to be factored into the design of autonomous ships. During the transition from conventional to autonomous ships, all past experience should be transferred to new systems. An anomaly detection system integrated with big data analysis, inference systems and cloud systems can become quite sensitive to maritime accidents.

The evolution of shipping is a gradual process, with progressive technological advances changing how decisions are made, actions initiated, and initiatives taken during navigation and maritime operations. Ships will evolve from human operated to fully autonomous vessels, authority will evolve from human to software and actions taken will evolve from human to systems.

Autonomous ship design and autonomy optimisation have been among the most popular research topics in the maritime literature in recent years [L1]. Automatic Anomaly Detection System (AADS) is an inevitable requirement for unmanned marine vehicles. For this reason, unmanned surface ships are operated either in places where risk is minimal or in experimental scenarios with managed malfunctions, dangers and accidents [L2].

The marine environment is uncertain, complex, and dynamic, with many parameters at play, and shipping operations involve hundreds of possibilities and risks. In addition, a value that is not determined as a risk for a situation can be very risky in another situation and

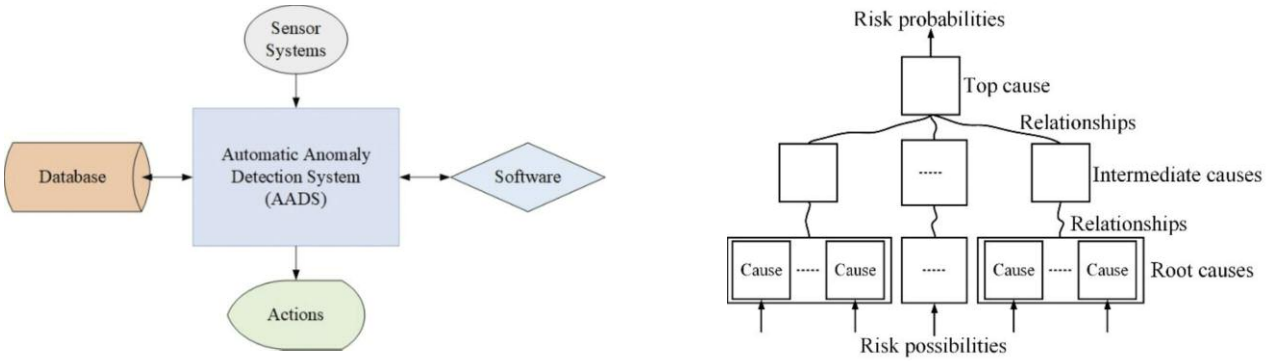


Figure 1: AADS model. Left (a) General framework of AADS for fully autonomous ships, right (b) Form of Risk Structures for AADS

vice versa. For example, an autonomous ship with a predetermined course is expected to deviate from the route against an obstacle and re-enter its route. But if this deviation coincides with the shallow water zone, then grounding can occur. Likewise, an action taken to avoid a known hazard can cause collision, grounding, machine failure, sinking or other accidents. The most important goal in the AADS is to determine the risk level for each action taken by the autonomous ship for all accident scenarios and recommend another action for risk values that are above the threshold values. The lack of a dynamic decision support system for unmanned marine vehicles threatens the safety of both vehicles and the marine environment. The AADS improves the safety of autonomous ships.

Autonomous ships should not only act according to the obstacles detected by the sensor systems, but also have the flexibility to include time, space, manual operation, and new rules. These rules can be determined in the light of past experiences, they can be taken automatically by cloud systems and adapted to the system with the help of artificial intelligence. Data entries can also be made manually. In the AADS model (Figure 1 (a)), inputs are obtained via sensor systems, the data and the risk structures of maritime accidents are collected in the database, and the data are processed via software.

AADS modifies and converts sensor data into risk possibilities. Risk possibilities are used to find the risk probabilities by using the risk structures. The structure of a risk model is given in Figure 1 (b). For example, a risk structure might be a collision, of which there are many different kinds. Therefore, the risk structures should be as flexible and comprehensive as possible. All possible scenarios are stored in the database to determine the accident risk at all times and locations. The mathematical logic behind the relationships requires an expertise and comprehensive analysis based on previous experience. Fuzzy systems for Bayesian networks, safety models such as fault tree analysis or T-S models are some of solutions for creating the risk structures [1-2]. The risk probabilities are the decision parameters of AADS. If the threshold is exceeded, immediate action is taken.

In the context of an ERCIM funded work, we focus on optimisation algorithms and quantitative decision support systems for maritime supply chains. We are currently working to improve the real-time continuous fuzzy fault tree analysis model [1].

Links:
[L1] https://kwz.me/h4z
[L2] https://kwz.me/h4A

References:
[1] Y. E. Senol, B. Sahin: "A novel real-time continuous fuzzy fault tree analysis (RC-FFTA) model for dynamic environment", Ocean Engineering, 127, 70-81, 2016.
[2] H. Pan, W. Yun: "Fault tree analysis with fuzzy gates", Computers & industrial engineering, 33(3-4), 569-572, 1997.

Please contact:
Bekir Sahin
NTNU, Norway
bekir.sahin@ntnu.no

Using Multiclass Classification for Ship Route Prediction

by Angelica Lo Duca and Andrea Marchetti (IIT-CNR)

The National Research Council in Pisa has been implementing a ship route prediction algorithm based on multiclass classification. The algorithm was developed within the OSIRIS project [1], which aimed to build a decision support system for maritime surveillance.

Multiclass classification (MC) is a type of supervised learning which measures the attributes of a sample, and based on those attributes, assigns the sample to one of many classes. MC is a generalisation of the binary classification, in which there are only two classes. MC can be exploited to build a ship route prediction (SRP) system [1], which aims to predict the next position of a ship, given its current status, determined by its current position (latitude, longitude), speed, direction, time and date, as well as a historical database of past routes. The area to be monitored,

called "area of interest" (AoI), is split into m rows and n columns, and each cell of the obtained matrix constitutes an output class of the MC algorithm. Thus, the total number of possible classes is m x n. Given the current status of a ship, the SRP system predicts the probability that each cell of the matrix will be occupied after a given period of time (e.g. 30, 45 or 60 minutes).

Different MC algorithms were tested to perform SRP, including naive Bayes, K-nearest neighbours (K-NN), decision trees, linear algorithms and extension from binary. Every MC algorithm was trained and tested with actual data from a historical database of past routes extracted from automatic identification system (AIS) messages sent by other ships around the island of Malta. A web application was then implemented to predict the next position of a ship [L2]. The K-NN and decision tree algorithms outperformed all the other MC algorithms.

The proposed SRP system is a point-based system, which only predicts the next position of a ship, given the current position. Other systems defined in the literature can predict the whole trajectory. To the best of our knowledge, our system is the first SRP algorithm to use time and date as input features.

We performed a qualitative test of the SRP system, which involved testing the SRP performance in the following real scenario: On 30 December 2019 at 11:11, the ship ASTRAEA [PA] was directed towards the Marsaxlokk port and the estimated time of arrival was 12:30. SRP was used to predict the next position after 60 minutes. All the information used for this scenario was extracted from Marine Traffic [L2], a global ship tracking intelligence website. We ran all the implemented MC algorithms to predict the ship's next position after 60 minutes. Results are shown in Figure 1: only K-NN and decision tree predicted the correct position.

The area around the island of Malta was used as the AoI, with periodic satellite images capturing the movement of ships within this region. Low-resolution was used for these initial images, to minimise costs. These images are processed by two modules, identification and classification, in order to extract the ships present in the area and their general status. The extracted status is sent as input to the SRP system, which identifies the position of the ship after a certain time interval. The position extracted from the SRP system is used to direct the satellite to capture a new image in the exact location that the ship is predicted to be found. The new captured image is more fine-grained to allow additional information to be extracted. This information will be used by another system, the behaviour analysis module, which tries to determine if the ships present in the subarea are behaving correctly or not. In summary, the OSIRIS system uses two satellite images, one coarse-grained (larger and less expensive), one fine-grained (smaller and more expensive). For this reason, the SRP system does not need to predict the exact position of a ship, but only a restricted area in which the ship will be after a given time interval.

Although SRP has been widely studied in the literature, this work is the first approach to compare different SRP algorithms based on multiclass classification. The use of machine learning techniques in the maritime field is being consolidated, thus helping decision support systems to profile, the then to identify anomalies, in ships' behaviours. The naive theoretical framework and apparently over-simplified assumptions of this methodology mean that it can easily be reproduced and used by other researchers to compare other families of algorithms, not limited to the field of SRP.

The SRP system was implemented within the OSIRIS project, funded by the European Space Agency (ESA) [1].

Links:

[L1] http://wafi.iit.cnr.it/osiris1/srp_viewer/
[L2] https://www.marinetraffic.com/

References:

[1] M. Reggiannini, et al.: "Remote Sensing for Maritime Prompt Monitoring", J. of Marine Science and Engineering, 7(7), 202, 2019
[2] A. Lo Duca, C. Bacciu, A. Marchetti: "A K-Nearest Neighbor Classifier for Ship Route Prediction", OCEANS 2017 - Aberdeen, UK 2017.

Please contact:

Angelica Lo Duca
IIT-CNR, Italy
angelica.loduca@iit.cnr.it

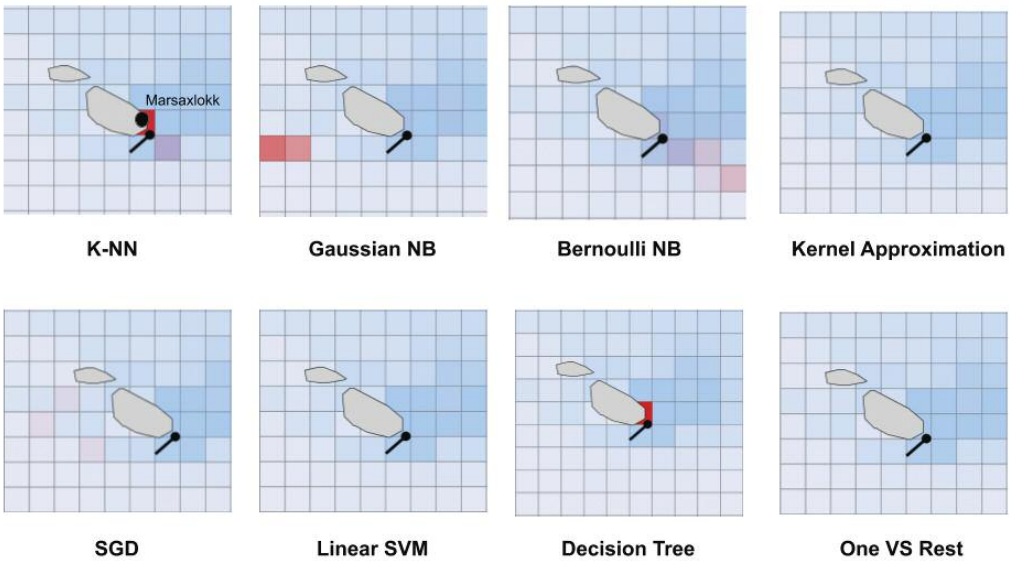


Figure 1: The output of SRP for the ship ASTRAEA [PA]. The position of the destination (Marsaxlokk port) is shown only for K-NN. Predicted probabilities are shown in red. The brighter the red of the cell, the higher the probability value.

ECAVI: A Teaching Assistant for Reasoning about Actions and Change

by Nena Basina, Theodore Patkos, Dimitris Plexousakis (FORTH-ICS)

Reasoning about actions, change and causality constitutes an important field of research in artificial intelligence. A visual representation of the main concepts involved while encoding logic programs can help knowledge engineers better understand the semantics. The ECAVI modelling tool aims to acquaint inexperienced modellers with the main features of common sense reasoning, guiding them during the process, through a meta-modelling platform and with the help of a state-of-the-art reasoner.

Action languages are well-established logical theories for reasoning about the dynamics of changing worlds. One of the most prominent, widely-applied action languages is the Event Calculus (EC), which incorporates useful features for representing causal and narrative information. EC implementations can be encoded in different languages, such as Answer Set Programming (ASP), a form of knowledge representation and reasoning paradigm oriented towards solving complex combinatorial search problems. ASP programs define a set of logical rules, whose models, called answer sets, correspond to solutions to a reasoning task, such as progression or planning. However, both ASP and the EC can be quite difficult to axiomatize by the non-expert and novice practitioners find it hard to properly model a domain of interest.

Some researchers [1], [2] have argued the importance of visualisations for knowledge engineers. In the context of EC, we argue that a visual representation of the various axiom types may help knowledge engineers understand the semantics of the different axiom types, thereby simplifying the learning process for inexperienced modellers and reducing the number of modelling mistakes.

To this end, we are developing a domain-independent tool, ECAVI (Event Calculus Analysis and Visualization) [L1], which offers a visual language for designing dynamic domains in the EC. ECAVI implements a translation of EC theories into ASP rules, through the ADOxx meta-modelling platform and with the help of the state-of-the-art automated Clingo ASP reasoner.

ECAVI is implemented with the use of the ADOxx meta-modelling platform [L3] where the developer can define the metamodel and the modelling method for the created toolkit. Following the framework for the description of modelling methods proposed by Karagiannis and Kühn [3], we developed a modelling language that is tailored to the Event Calculus way of representing causal relations. The ADOxx platform comprises two components: the Development Toolkit, which we use to build the modelling language of the tool, and define its graphical representation; and the

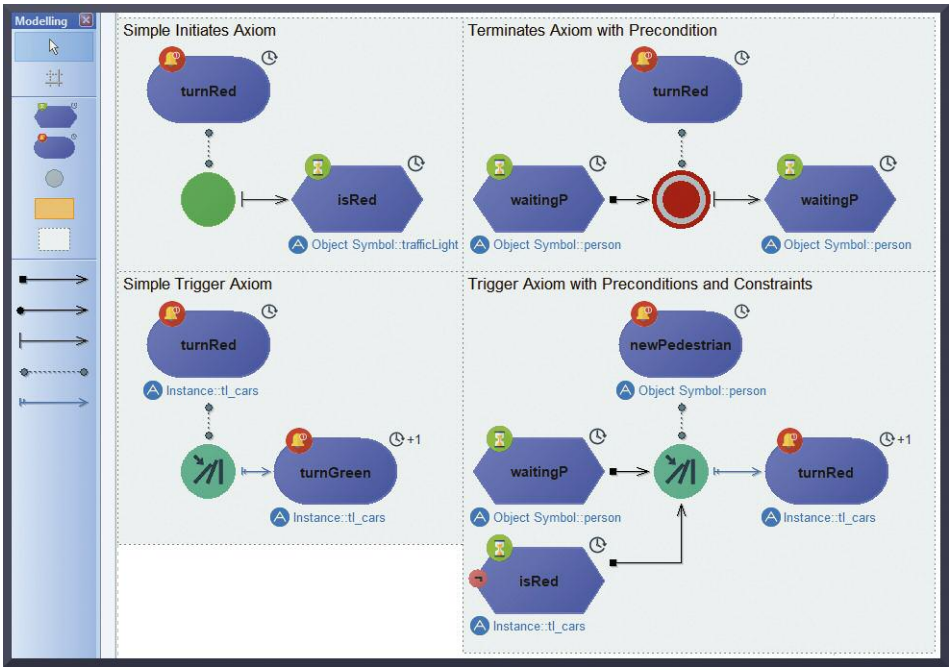


Figure 1: A sample of the axioms defined for the traffic light paradigm.

Modelling Toolkit, where the end user designs a domain of application that will be translated into an ASP program. The program is then sent to the Clingo ASP reasoner [L2] to produce the results. The AdoScript macro language of ADOxx is the actual link between ADOxx and Clingo, while a Java program is used as an intermediate for translating the designed models into ASP and vice versa. To accommodate the modelling process, we follow a common practice in knowledge engineering for dynamic domains, which breaks down the modelling tasks into four sub-models, separating the "alphabet" from the domain axiomatization.

We also use AdoScript to implement a number of features that assist the user during the process of building an axiom, to support a number of integrity checks and for implementing simple, yet fundamental checks for syntactical errors. In the future, we aim to implement some extensive meta-reasoning into ECAVI that will help more experienced users better run and visualise their programs.

The target group of ECAVI is novice knowledge engineers, with little or no knowledge of the EC formalism. To facilitate the use of the tool, a tutorial has been implemented, using traffic lights as a sample use case, as part of a general-purpose smart city scenario. The tutorial is in the form of a walk-through wizard, which guides the user step-by-step through the process of defining the domain of interest. A snapshot of the axioms that define the use case is shown in Figure 1.

Overall, ECAVI aims to offer a visual language for designing causal dynamic domains, achieving a tight coupling of the visual domain representation with two powerful logical formalisms, namely the EC and ASP, while also assisting the user in the process of knowledge engineering, minimising the syntactical errors and helping with logical fallacies. And most importantly, ECAVI adopts a pedagogical approach and aims to help the non-expert learn the basics of how a conceptual model can be translated and executed through the logic programming paradigm.

ECAVI will be extended in the future. We are currently working on evaluating the tool's usability in terms of satisfaction, efficiency and effectiveness. Upcoming versions will consider more features such as non-determinism, indirect effects of events, causal constraints, etc. Furthermore, we plan to extend our focus to other types of users, with different levels of modelling experience. As more experienced users have different needs, new features will have to be supported, which may be useful for modelling large domains with complex rules. In the long run, we envision ECAVI to take the form of a fully visual integrated development environment for modelling dynamic domains, complemented with a debugger, step-by-step execution and other features typically found in IDEs.

Links:

[L1] https://nenabas.github.io/ECAVI_release/
[L2] https://potassco.org/
[L3] https://www.adoxx.org/live/introduction-to-adoxx

References:

[1] R. Morgan, G. Grossmann, M. Stumptner: "VizDSL: Towards a Graphical Visualisation Language for Enterprise Systems Interoperability", 2017 International Symposium on Big Data Visual Analytics (BDVA), 1–8, 2017
[2] H-G Fill, D. Karagiannis: "On the conceptualisation of modelling methods using the ADOxx meta modelling platform", Enterprise Modelling and Information Systems Architectures, 8:4-25, 2013.
[3] D. Karagiannis, H. Kühn: "Metamodelling platforms", in Proc. of EC-WEB '02, pp. 182-, Springer-Verlag, 2002.

Please contact:

Nena Basina
FORTH-ICS, Greece
basina@ics.forth.gr

Observing Taxi Behaviour at Charging Stations and Taxi Stands Using Image Recognition

by Maarten Groen (Amsterdam University of Applied Sciences) and Nanda Piersma (Amsterdam University of Applied Sciences, CWI)

City authorities want to know how to match the charging infrastructures for electric vehicles with the demand. Using camera recognition algorithms from artificial intelligence we investigated the behavior of taxis at a charging stations and a taxi stand.

In Amsterdam, the municipality has placed fast charging stations throughout the city to support electric taxi ownership. This represents a step towards its ambitious goals, which include prohibiting all vehicles except "green taxis" from using taxi stands in prime city locations, such as train stations and inner-city tourist attractions, and even entire city areas.

Because of the proactive attitude of all stakeholders, increasing numbers of charging stations are being established throughout the city. When a taxi is connected to a charging pole, both the transaction for each car and the usage of each station is registered in a central database. However, since information is only collected when vehicles are connected to stations, we lack data about drivers that fail to charge their vehicle when all stations are occupied. Additionally, it is unclear how often taxi-stands are still being used by customers with the increased availability of taxi-apps. These datapoints are the missing link for policy making on charging infrastructure demand and the usage of taxi stands.

In a joint project with the Municipality of Amsterdam, the use of charging stations and taxi stands is monitored to better understand charging demand and taxi-stand customer activity. This information will help the municipality in getting to more informed policy decisions for their ambitious 'green taxi' plan. With an image recognition algorithm, we observed a fast charging station and a taxi stand in Amsterdam, the Netherlands and counted events related to taxis. The automation of the observation was found to be non-trivial.

The challenge

Ideally, for the sake of privacy, camera footage should not be stored when it may contain images of individuals such as taxi drivers or members of the public and when tracking of movement of found objects is required. For this reason, the speed of the algorithm to process the camera footage so that no data has to be stored was an important consideration when developing the algorithm.

State-of-the-art object detection algorithms have been rapidly improving in accuracy and speed [R1]. It is now easy to find cars, trees, people or charging stations (after some training) within an image. In this project the algorithm used object detection and object tracking ([R2],[R3]) as the major building-blocks to be able to distinguish between: taxis and other vehicles using the charging stations, cars waiting for a charging station, taxis waiting for customers or taxis taking a break, and the interpretation of drive-by cars leaving because the charge stations are full. Finally, the algorithm should use as many generally available trained elements as possible to facilitate transfer of the algorithm to other locations without extra location-specific training.

Results

We collected and analyzed data at two locations, a fast charging location and a taxi pickup location. The fast charging location, shown in Figure 1, opened in May 2019 with four fast charging stations. The location can only be reached through a dedicated road with a dead end. Figure 2 represents an overview of the taxi stand at the Amsterdam main train station, showing four places where taxis can pick

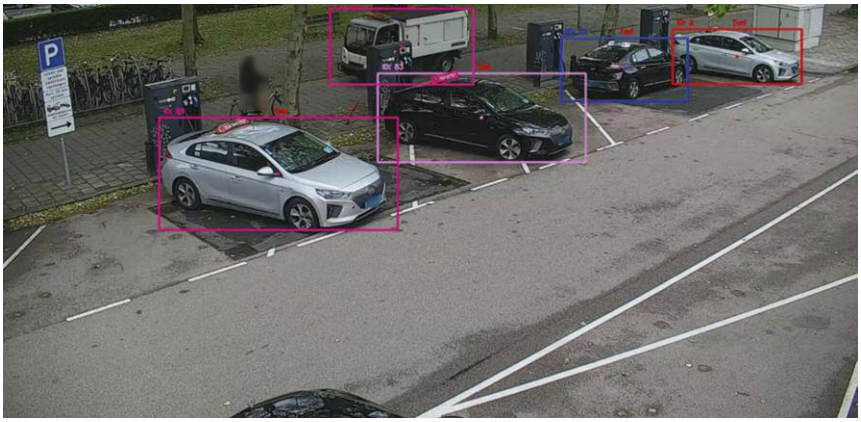


Figure 1: Example from data collected at the fast charging location.

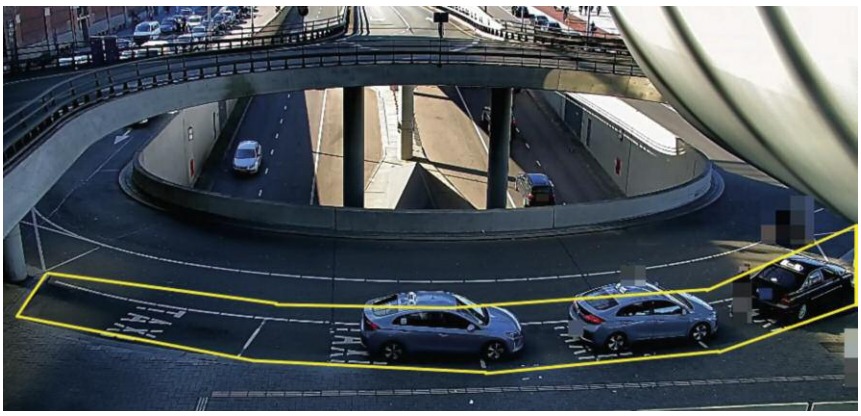


Figure 2: Example from data collected at the taxi stand.

up customers. Taxis are directed to the pickup location with an automated number plate-based system. We focused on the pickup location to count the taxis leaving with customers (versus taxis leaving without customers).

We successfully identified car types (taxi or other), charging taxis, taxis waiting to charge and

taxis driving away. This enabled us to count taxis using the fast charging stations, with duration (time) slots for all activities. The taxi stand was much busier than the fast charging station, with many pedestrians and cars not related to the taxi pick up process moving through the video. We could identify and count the number of taxis picking up customers (per time slot). Owing to the many passers-by in the vicinity of the taxis, the algorithm could not be trained to identify the number of customers entering a taxi. The object tracking task was especially challenging due to unexpected walking patterns at this location; people passing the waiting taxis and leaving the images were often recorded as entering the taxi. In addition, owing both to the camera position and the fact that taxi drivers often leave the car multiple times, drivers are difficult to distinguish from customers.

The simpler tasks can be done in real time, taking an average of 13 minutes to analyze a 15-minute video. This includes tasks such as identifying and counting taxis and determining time slots. More complex tasks, such as counting the number of people entering a car, require more computational power and more location-specific training to achieve acceptable results.

The Municipality of Amsterdam is considering applying the new algorithms to new taxi stands and to other use-cases. This research is part of the research on energy transition of the Intelligence and Autonomous Systems group of the CWI and the IDOLAAD project [L1] at the Amsterdam University of Applied Science. Future research will focus on exploring ways to further automate the more complex tasks.

Link:
[L1]: https://www.idolaad.com/research/research.html

References:

[1] J. Redmon, A. Farhadi: "Yolov3: An incremental improvement", arXiv preprint arXiv:1804.02767, 2018.
[2] L. Leal-Taixé, et al.: "Tracking the trackers: an analysis of the state of the art in multiple object tracking", arXiv preprint arXiv:1704.02781, 2017
[3] S. R. E. Datondji, et al.: "A survey of vision-based traffic monitoring of road intersections", IEEE transactions on intelligent transportation systems, 17(10), 2681-2698, 2016

Please contact:

Maarten Groen, Amsterdam University of Applied Science, Netherlands, m.n.groen@hva.nl

Nanda Piersma, Amsterdam University of Applied Science, CWI, Netherlands, nanda.piersma@cwi.nl

Securing Home Automation Systems against Sensor Manipulation

by Albert Treytl, Edith Huber, Thilo Sauter (Danube University Krems) and Peter Kieseberg (St. Pölten University of Applied Sciences)

Home automation systems (HAS) can be important attack vectors, yet research on securing sensors is sparse, especially with respect to the analogous side of these components, i.e., detecting manipulations of the sensors themselves. Metadata together with the combination of several sensor nodes can be used to thwart such manipulation attacks.

The Internet of Things is a wide and diverse ecosystem where interconnected devices and services collect, exchange and process data in order to adapt dynamically to a context, automate tasks and provide better situation awareness to react to customer needs. Home automation systems (HAS), which are commonly based on IoT, are a growing field for many applications, such as comfort, surveillance and access and energy saving. Since HAS generate a lot of private data, they are very appealing to attackers, who can use them to spy on or stalk inhabitants, or use them to facilitate more traditional criminal activities like burglaries [1]. The comprehensive interconnection of systems to an Internet of Things offers enormous potential to HAS, but also generates new cyber-risks. This has been discussed in-depth by many other researchers, often in the context of industrial or workplace environments, such as building automation and industrial IoT, pointing out that the quality and/or veracity of the source information, typically provided through sensors, forms the basis for securing IoT systems. Thus, both the acquisition and the communication of this information requires special attention in an IoT-environment.

Most of the basic HAS standards currently in use were developed the from the late eighties to early nineties, and IT-security, such as KNX, was added later on. There remain many open questions and challenges, especially in relation to security measures that rely directly on the sensor data and related meta-information. While there are several approaches to use meta-information to discover malicious software (e.g., [2]), the analogue side of the sensors (hardware) is typically neglected, even though manipulation on this side makes typical countermeasures obsolete. This also applies to the extraction of meta-information in the analogue sensor circuit, which could help detect such manipulations and thus help close the security gap in sensor systems [3]. Thus, two different attacker approaches must be considered for home automation systems:

- An attacker could manipulate the data in the digital realm, i.e., the sensor sends correct information, but it is modified in the network. This typical approach is often referred to in the academic literature. Even in this context new tech-

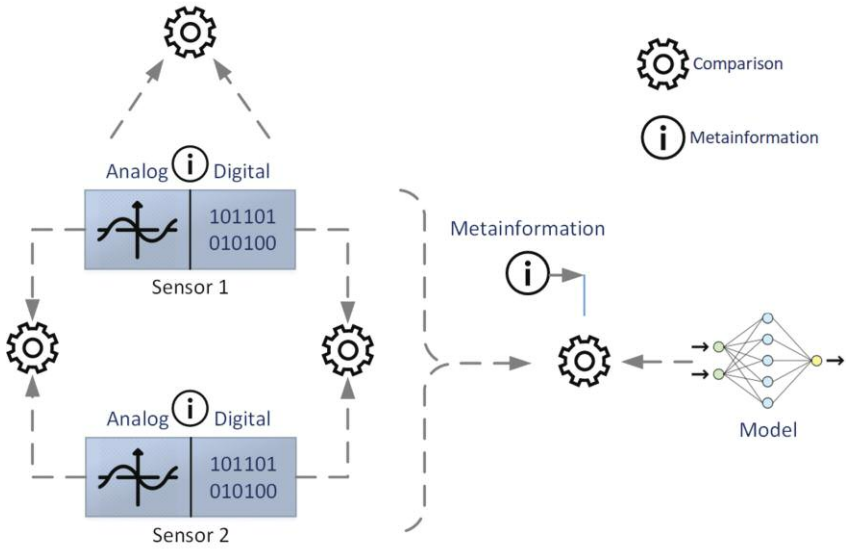


Figure 1: Comparing analogue and digital information and meta-information.

niques are required owing to the low performance of cheap sensors that cannot support standard techniques like signatures or other cryptographic schemes.

- Alternatively, the attacker could manipulate the sensor itself to make it publish incorrect information to the network. Particularly at the analogue level, substantial research is required on both sides: from the perspective of the attacker, to better understand the possibilities in real-world examples; and from the perspective of defence, to better understand these attacks and their impacts on meta-information in order to provide new means of defence.

While the effects of these attacks look similar, the mitigation strategies required to overcome them are very different and require different sets of skills, and the analogue side or circuit has been largely neglected in the security community. When considering meta-information, attacks could potentially be made much harder when extending the view of the defender from the single sensor to related meta-information and further to a multi sensor scenario, where information and especially meta-information, needs to fit together, i.e., the attacker needs to coordinate different attacks to remain undiscovered in a more complex physical model.

In our project "ARES", we will therefore analyse the extended attack surface directly on the analogue side of common sensors, considering payload information as well as meta-information. A combination of sensor data and highly attack resistant meta-information will be used to increase the overall attack resistance. Through the combination of information from multiple sensors, as well as matching with (physical) models, the attacker would need to stay consistent between manipulated sensor signal and meta-information within a single device but also over several devices, making the attack much more difficult to carry out and thus raising the bar (see Figure 1).

Additionally, the success of technical measures always depends on user acceptance. International organisations, such as EUROPOL and ENISA, regard a multidisciplinary approach as indispensable, since HAS suffer from a security hostile environment including low cybersecurity awareness of users, as well as fast and extremely low cost implementation and unplanned installation compared to other fields such as industrial automation [1]. The continuous expansion of digitalisation into daily domestic life raises questions that are best addressed from a social science perspective, incorporating sociology, legal studies and economics. ARES will investigate these issues, providing an evidence-based analysis of cyber-risks and user requirements, and will use these findings to improve the design of the security measures.

References:

[1] E. Casey: "Digital evidence and computer crime: Forensic science, computers, and the internet", Academic press, 2011

[2] A. Sadighian, et al.: "A context-aware malware detection based on low-level hardware indicators as a last line of defense", 2017.

[3] P. Palensky, T. Sauter: "Security considerations for FAN-Internet connections" in 2000 IEEE Int. Workshop on Factory Communication Systems, Proceedings, Cat. No. 00TH8531, pp. 27-35, IEEE 2000.

Please contact:

Peter Kieseberg
University of Applied Sciences, St. Pölten, Austria
peter.kieseberg@fhstp.ac.at

SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Call for Proposals

Dagstuhl Seminars and Perspectives Workshops

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is accepting proposals for scientific seminars/workshops in all areas of computer science, in particular also in connection with other fields.

If accepted the event will be hosted in the seclusion of Dagstuhl's well known, own, dedicated facilities in Wadern on the western fringe of Germany. Moreover, the Dagstuhl office will assume most of the organisational/ administrative work, and the Dagstuhl scientific staff will support the organizers in preparing, running, and documenting the event. Thanks to subsidies the costs are very low for participants.

Dagstuhl events are typically proposed by a group of three to four outstanding researchers of different affiliations. This organizer team should represent a range of research communities and reflect Dagstuhl's international orientation. More information, in particular, details about event form and setup as well as the proposal form and the proposing process can be found on

https://www.dagstuhl.de/dsproposal

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is funded by the German federal and state government. It pursues a mission of furthering world class research in computer science by facilitating communication and interaction between researchers.

Important Dates

- Proposal submission: July 1 to July 15, 2020
- Notification: October 2020
- Seminar dates: Between May 2021 and August 2022 (tentative).

Please note, there will be no submission deadline in November. The next following submission deadline will be April 15, 2021.

ERCIM "Alain Bensoussan" Fellowship Programme

The ERCIM PhD Fellowship Programme has been established as one of the premier activities of ERCIM. The programme is open to young researchers from all over the world. It focuses on a broad range of fields in Computer Science and Applied Mathematics.

The fellowship scheme also helps young scientists to improve their knowledge of European research structures and networks and to gain more insight into the working conditions of leading European research institutions. The fellowships are of 12 months duration (with a possible extension), spent in one of the ERCIM member institutes. Fellows can apply for second year in a different institute.

Why to apply for an ERCIM Fellowship?

The Fellowship Programme enables bright young scientists from all over the world to work on a challenging problem as fellows of leading European research centers. In addition, an ERCIM fellowship helps widen and intensify the network of personal relations and understanding among scientists.

The programme offers the opportunity to ERCIM fellows:

- to work with internationally recognized experts;
- to improve knowledge about European research structures and networks;
- to become familiarized with working conditions in leading European research centres;
- to promote cross-fertilization and cooperation, through the fellowships, between research groups working in similar areas in different laboratories.

Conditions

Candidates must:

- have obtained a PhD degree during the last eight years (prior to the year of the application deadline) or be in the last year of the thesis work;
- be fluent in English;
- have completed their PhD before starting the grant.

The fellows are appointed either by a stipend (an agreement for a research training programme) or a working contract. The type of contract and the monthly allowance/salary depends on the hosting institute.

Application deadlines

Deadlines for applications are currently 30 April and 30 September each year.

Since its inception in 1991, over 500 fellows have passed through the programme. In 2019, 53 young scientists commenced an ERCIM PhD fellowship and 79 fellows have been hosted during the year. Since 2005, the Fellowship Programme is named in honour of Alain Bensoussan, former president of Inria, one of the three ERCIM founding institutes.

http://fellowship.ercim.eu

ERCIM
European Research Consortium for Informatics and Mathematics

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Host of the World Wide Web Consortium.

CNR
Consiglio Nazionale delle Ricerche
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1, 56124 Pisa, Italy
www.iit.cnr.it

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

CWI
Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
www.cwi.nl

RISE
RISE SICS
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Fonds National de la Recherche Luxembourg
Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
www.fnr.lu

SBA Research
SBA Research gGmbH
Floragasse 7, 1040 Wien, Austria
www.sba-research.org/

simula
SIMULA
PO Box 134
1325 Lysaker, Norway
www.simula.no

FORTH
INSTITUTE OF COMPUTER SCIENCE
Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
www.ics.forth.gr

MTA SZTAKI
Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
www.sztaki.hu/

Fraunhofer
IUK-TECHNOLOGIE
Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
www.iuk.fraunhofer.de

TNO innovation for life
TNO
PO Box 96829
2509 JE DEN HAAG
www.tno.nl

inesc
INESC
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal
www.inesc.pt

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
www.cs.ucy.ac.cy/

Inria
Institut National de Recherche en Informatique et en Automatique
B.P. 105, F-78153 Le Chesnay, France
www.inria.fr

UNIVERSITAS VARSOVIENSIS
Universty of Warsaw
Faculty of Mathematics, Informatics and Mechanics
Banacha 2, 02-097 Warsaw, Poland
www.mimuw.edu.pl/

I.S.I.
Industrial Systems Institute
I.S.I. – Industrial Systems Institute
Patras Science Park building
Platani, Patras, Greece, GR-26504
www.isi.gr

VTT
VTT Technical Research Centre of Finland Ltd
PO Box 1000
FIN-02044 VTT, Finland
www.vttresearch.com

Figure 1: