



Cloudscape Brazil 2018

Trusted Technologies for strong and competitive economies



Workshop on Cloud Networks

ATMOSPHERE

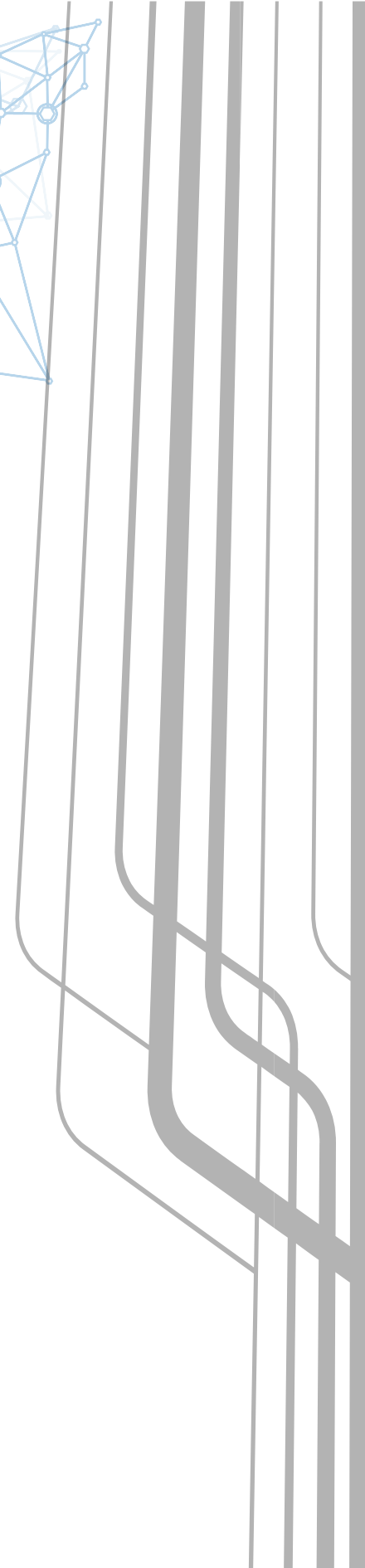
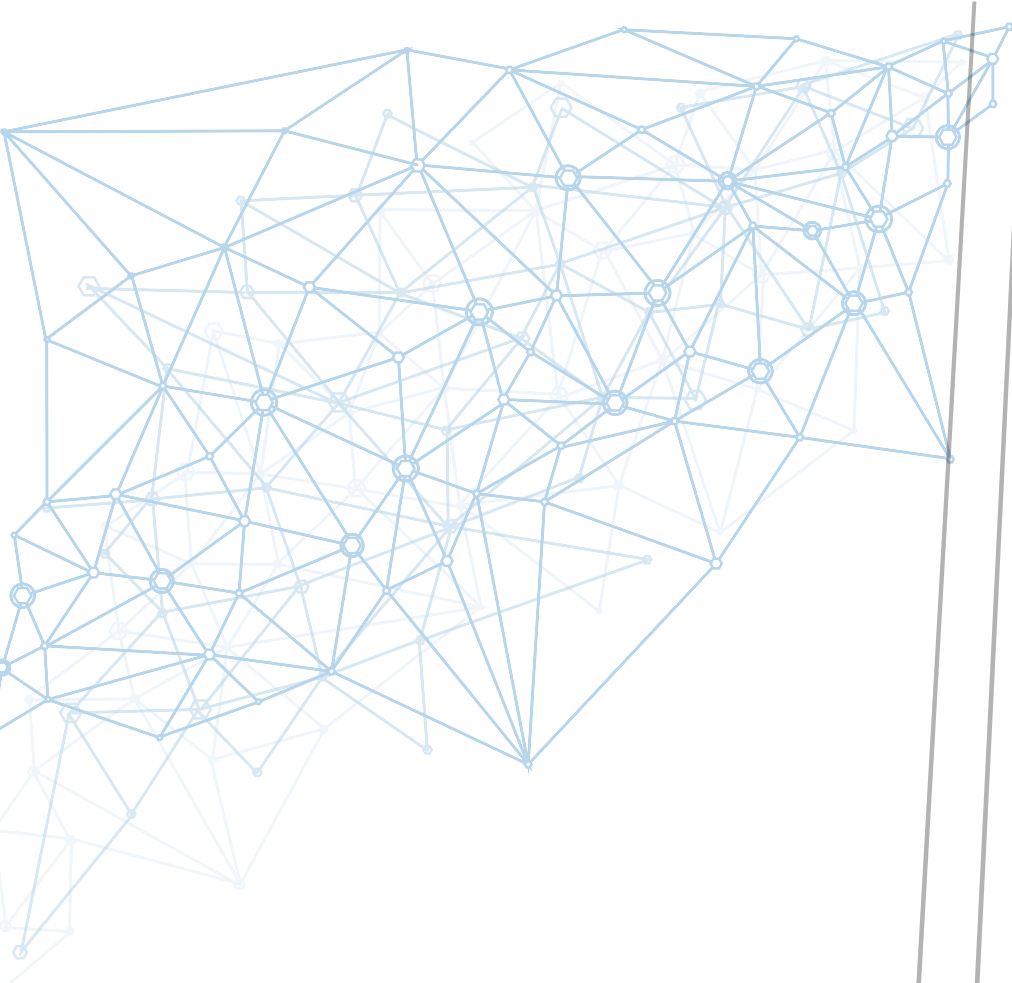
Adaptive, Trustworthy, Manageable, Orchestrated, Secure Privacy-assuring Hybrid, Ecosystem for RESilient Cloud Computing



EUBrasilCloudFORUM

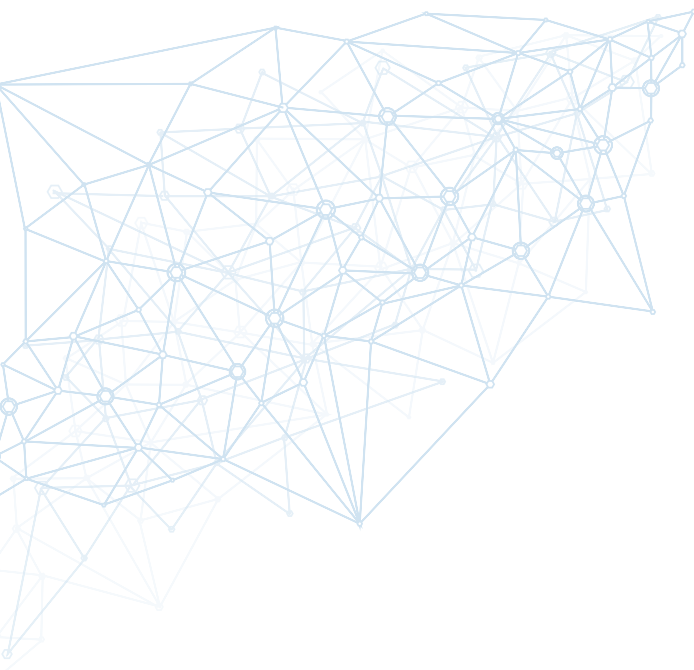
Fostering an International dialogue between Europe & Brazil

POSITION PAPERS



Index

- Welcome message 3
- Artificial Intelligence 4
 - Semantic Analytics: The accelerator of Artificial Intelligence Digital Markets..... 5
 - A Rheumatic Heart Disease Classifier Based on Echocardiographic Data 8
 - Artificial Intelligence for Scheduling Resource Blocks in LTE/5G Networks 9
- Competitive European/Brazilian SMEs 11
 - Cloud-based time-limited transaction management 12
 - Support SMEs in becoming competitive and exploiting the potential of international markets: EU-Brazil SME coo..... 14
- Digitizing Industry 16
 - The Impact of 5G on Vertical Industry Sectors..... 17
 - FASTEN – Flexible and Autonomous Manufacturing Systems for Custom-Designed Products..... 19
- EU-Brazil Common standards 22
 - A Proposal to Apply a Risk Assessment Methodology for IoT Systems to a Smart Childhood Obesity Caring Solution 23
 - Protecting sensitive data in cloud environments 26
- International Partnerships 28
 - What will the future hold for EU-BR collaboration in ICT? 29
- Regulation for Trust in Digital Environments 32
 - How much can I trust my cloud services?..... 33
 - Fairness and Transparency in Trustworthy Cloud-based Analytics Services..... 35
 - Device-Based Security to Improve User Privacy in the Internet of Things..... 37



Welcome message

We are proud to share with you the proceedings of the fifth edition of Cloudscape Brazil and third edition of Workshop on Cloud Networks (WCN), both events organised within the remit of the ATMOSPHERE project, support by EUBrasilCloudFORUM project, which organised the previous two editions.

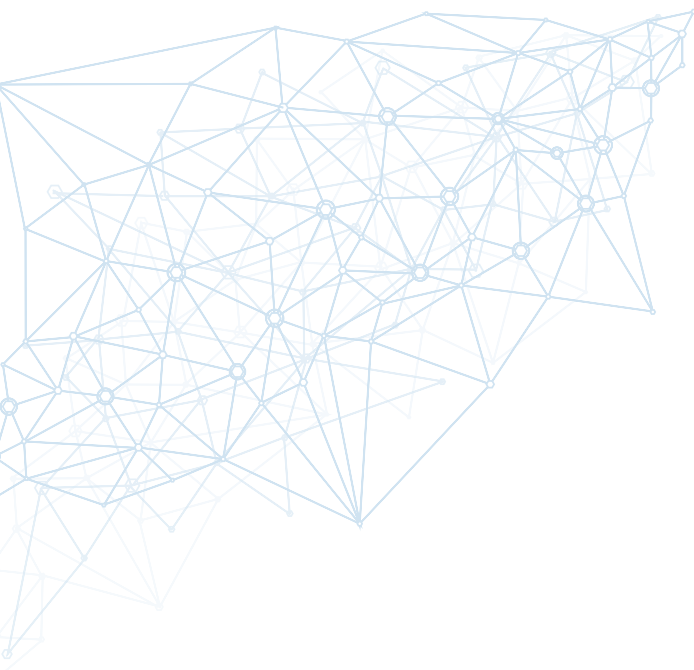
Cloudscape Brazil took place on 25 July, while WCN on 26 July, in Natal (Brazil), co-located with the annual Congress of the Brazilian Computing Society (CSBC), for the third time.

We would like to thank the Congress of the Brazilian Computing Society for being our hosts, to Brasscom for the institutional support and to all those who submitted Position Papers this year.

We wish all of you an interesting reading.

For more information about Cloudscape Brazil & WCN events, please visit:

<https://www.atmosphere-eubrazil.eu/cloudscape-brazil-and-workshop-cloud-networks-2018>



Artificial Intelligence



Semantic Analytics: The accelerator of Artificial Intelligence Digital Markets

Author: Dr. Martin Serrano (Insight Centre for Data Analytics Galway)

Who stands to benefit and how

General Audience, experts in Analytics tools and expert on Interpretation about the findings of the data after running an analytics process. Also, in the recent hype for Artificial Intelligence (AI), Data experts and specialist on Artificial Intelligence can benefit from this position paper information, first to situate the current state of the art and second to understand the emergence and use of Semantic Analytics towards enhancing/accelerating Artificial intelligence

Position Paper

Data is everything today, everywhere and everybody can produce, process and consume data, from a simple sensor detecting our presence and an actuator providing a message or activating a door to be opened to a more complex system running expert computing like in economy or stocks markets where the data defines the rise or down of the indexes. But what are the elements that make the data so useful and rich? The answer is simple: "Analysis and Interpretation" The phenomenon of everybody producing data continues growing and especially with the growth of Internet of Things technology (IoT), but volume is not all, else to make these data useful. It is necessary to provide meaning and extract information in the data and most of the time transform it into information and then eventually in knowledge enabling intelligence. Internet of Things [1] is expected to reach every aspect of our human activity and similarly cloud computing is now immersed and everybody use it indistinctly if it is for only storage in cloud or larger performance using the cloud it is expected to reach more to virtual activity too increasing exponentially the volume of information, thus it is important to identify clearly the analytics and their tools.

The big diversity on analysis and analytics tools has made the Analytics area so rich in vocabulary that today it is even difficult to differentiate if there is one or multiple type of analytics. Analytics today is a capacity, a feature to understand data and the best way to materialize value (including economic) to the data but what Analytics is about? what makes analytics so different from one to another domain? and why Analytics methodologies/technology cannot be used equally in every domain? and particular now when Artificial Intelligence (AI) is hype, how analytics will improve the performance of AI systems. To resolve these questions let's first understand there are analytics tools that are used for different purposes.

The first one is those analytics tools prepared for use data coming from physical devices (also call raw data) like IoT sensors, the second ones are those analytics tools which are web-based (which means that the tool is using online data, commonly refers to analytics interface or analytics API) and the third one those analytics tools for the web which analyse specific structured data in order to improve web user experience. Having those different use cases using processing data the last one is visualising data, the ultimate outcome of the analytics is visualisation and interpretation.



Analytics

The majority of analytics tools are not considered suitable for a single (or several) specific domain(s), which makes difficult to define the best analytic tools, if this does not rely in a particular form of data or a particular domain of application (e.g. stock markets, production, values and offers, etc) . It is not possible that analytics tools seamlessly adapt for multiple domains. For example, we can take example of IoT data analytics like Kaa [4], Kaa can be used in a lot of IoT use cases, and for each one Kaa give some examples of use. However, even if the majority of tools are designed for IoT use cases, very of them cannot be used for other than collect and process values in that particular domain. Other analytics tools have been extended to applications and aiming to be generic, the first example like this is AWS-IoT Analytics [2], and even at AWS IoT there are use cases which are: the smart agriculture, the predictive maintenance, the proactive replenishing of supplies and the process efficiency scoring, the objective is to specialize the operation and provide meaningful interpretation about the data. Another tool seen before with specific use cases is Thingsboard [5], Thingsboard is designed for four use cases: the smart metering, the smart energy, the smart farming and the fleet tracking (bus tracking in the demo). Taken just as example the IoT domain it can be understood that analytics and its tools are domain dependent and thus the more specialised tool to a particular domain data the better are the analytics results.

Artificial Intelligence

Considering the analytics tools for IoT, let's keep in mind the same domain of application but now let's consider those analytics tools (for any types of data) which are web-based (online). It is important to know that some analytics tools are also platforms which work not only online and then they can be used to analyse website traffic like the users' clicks or the time spent by the user on the site for example. The most used AI tools today for IoT are Microsoft Azure Stream Analytics, AWS IoT Analytics, SAP Analytics Cloud and IBM Watson IoT Platform. Azure [1], can be integrated with two other tools to make better analytics (in real time for example). The AWS tool is a tool created by Amazon, its advantages is the free trial period during 12 months which allows to practice and understand this platform before to pay [2]. When we talk about Artificial Intelligence (AI) we are talking a totally computer-based technique from the collection of data, processing to analysis, interpretation and visualisation, the human intervention is minimised and for this reason AI is many times mixed with machine learning as a technique, but the main difference is that AI main outcome is knowledge generation while machine learning focuses in enhancement of the data and interpretation.

Semantics

Semantics is well known in the domain of web services, everybody has used a search engine or take the benefits from an indexed structured data base, now imagine all this benefits using a distributed world-wide data base, the benefits are exponential and highly evident, what makes powerful this approach is largely the distribution and the infrastructure but most important is the structure of the data, think in the way you can understand a language when you are traveling from one side to other side of the world, different accents, different interpretation and sometimes different meaning but when the structure exist it is much easier to understand and interpret the meaning of a sentence. In the same example, when you have relevant and useful information and the structured data and at the same time you can find an answer to a question without having to ask a human the benefits exponentially increase, thus the meaning of

colloquial speech, using semantics is all advantages from observing uncover specific meanings of words used in foreign languages mixed with our own to get more into real meanings. This position paper triggers the discussion on how analysis and their analytics tools, can be used in a particular data domain i.e. IoT, and how every types of data with requires specialised tools for better results when analysis is required. Also, the web-based analytics tools and analytics tools for web are mention as the way to enhance Artificial Intelligence results.

A Rheumatic Heart Disease Classifier Based on Echocardiographic Data

Authors: André Meirelles (UnB), Willian Barreiros (UnB), Alba Melo (UnB), Wagner Meira (UFMG), Bruno Nascimento (UFMG), Maria C. Nunes (UFMG), Antonio Ribeiro (UFMG), George Teodoro (UnB)

Who stands to benefit and how

Research and industry groups working with artificial neural networks can use this work as example to encourage the use of cloud-based environments as development and production tools with medical applications. Furthermore, this work is part of the ATMOSPHERE project, and will be employed to validate its use.

Position Paper

Convolutional Neural Networks (CNNs) have already been established as an efficient tool for a number of image processing tasks, such as, image classification, object identification, and face recognition. Through the use of CNNs, in this work, multiple CNN models will be proposed and evaluated according to their appropriateness to the interpretation of echocardiograms. The focus of the interpretation is on a specific heart condition, known as the rheumatic heart disease (RHD), which is a leading cause of heart failure in unprivileged populations.

8

The RHD is a heart condition caused by abnormal immune response to streptococcal infection, which is a bacteria normally associated with poor sanitation and hygiene conditions. The disease mainly attacks the mitral and aortic valves, resulting in permanent disability or death on the worst case, if not properly treated. Clinical diagnosis of RHD is dependent on auscultation and detection of valvular murmur, which is only possible when valve damage is more significant. Echocardiographic exams are an important tool for treating RHD, because it enables early detection of the disease and adequate treatment.

Furthermore, RHD is the highest worldwide cause of morbidity and mortality among heart valve diseases. One of the reasons for such is that effective detection of early RHD requires echocardiographic imaging, which can be inaccessible for certain regions (low-income and/or rural areas). As a solution, the UFMG PROVAVAR group proposed the use of hand-held echocardiographic devices by non-experts as a way to improve the accessibility of RHD diagnoses. Nevertheless, this solution requires a large amount of human resources. These non-experts also need to undergo a training process that enables them to recognize RHD features in echocardiograms.

In this work, we propose to improve the PROVAVAR approach leveraging the artificial neural network to improve the accuracy of RHD identification, while also reducing the human cost, on top of a cloud-based environment to ensure availability and data security. The use of the ATMOSPHERE environment will be essential for such endeavor, providing the trustworthiness and availability requirements for this medical application. Also, it will facilitate the development and improvement of the neural network itself, easing the optimization process for the network hyper-parameters.

Artificial Intelligence for Scheduling Resource Blocks in LTE/5G Networks

Author: Guilherme Branco (UnB)

Co-authors: Gabriel Carvalho (UnB), Marcos Caetano, Priscila Solis (UnB)

Who stands to benefit and how

Research groups that are working with 5G Projects

Position Paper

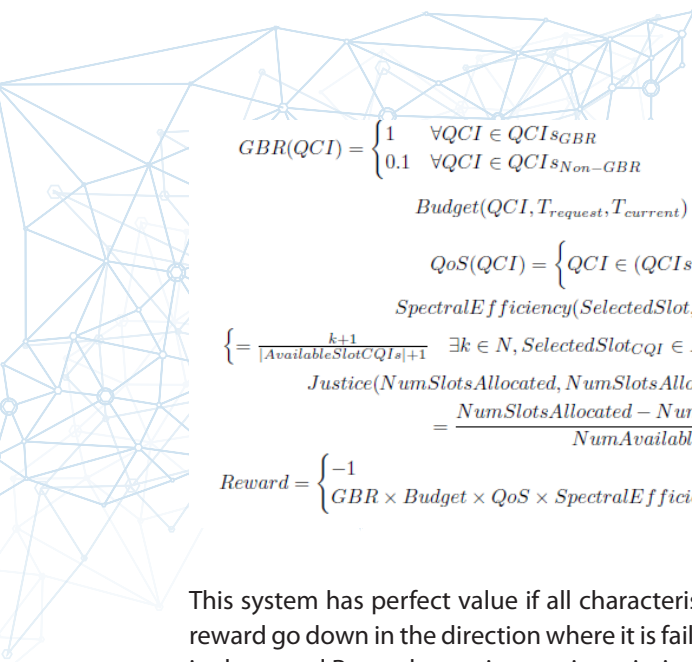
Connected devices have wildly different traffic patterns, and with 5G these patterns are expected to become even more different due to better infrastructure allowing support of different latencies, throughput and other Quality of Service (QoS) requirements. In such a scenario, the already crowded spectrum will end up excessively populated as there will be far more devices competing for transmission spectrum, on those grounds resource allocation needs to guarantee a minimum level of service for users in crowded places.

Resource scheduling is a complicated problem due to multiple goal optimization that are mutually exclusive, including, but not limited to, latency, throughput, fairness and spectrum efficiency [1]. A lot of heuristics are already in use to solve and improve block allocation, one of the simplest and most naive solution is Round-Robin scheduler, and a bit more complex ones with Genetic Algorithms (GA) [2] or other techniques.

Optimal solution would require knowledge of the future to allow scheduling of resources in optimal manner, that is impossible to be done in real-time, and even if would be possible due to the sheer amount of data there would need to be a tremendous amount of processing power to solve the problem in the short amount of time imposed by mobile networks standards. One possible solution comes by utilizing Artificial Neural Networks (ANN) to find a sub optimal solution training with real traffic data.

The 5G-Range Project expects 5G networks to reach 50Km range of coverage with focus in rural areas, where traffic is far different in comparison to urban areas. Moreover in rural areas there is an abundance of white-spaces in the spectrum, that is caused by lower demand over the licensed channels as well as in the unlicensed bands, making a huge improvement in spectrum efficiency if those holes can be used for something positive, such as Device to Device (D2D) or Machine to Machine (M2M) communications, it can also be used for traffic offloading from mobile network to unlicensed band through Wi-Fi or any other wireless technology from that band. On these facts a traffic generator, which is based on QoS Class Identifier (QCI) values, that range from 0 to 15 and can be used to classify the type of application [3].

For mobile networks, an ANN with unsupervised learning can be used with reinforcement learning, where a fitness value or reward is given to each action, prediction, done by the ANN, by doing that operation several times and going through the data. In different epochs, the neural network is able to learn fairly well how it should allocated resource blocks based on how much reward it is getting as feedback [4]. Manipulating the reward system more weight can be given for a specific characteristic that one might need to optimize, such as spectrum efficiency or latency. Current reward system is as follows:



$$\begin{aligned}
 GBR(QCI) &= \begin{cases} 1 & \forall QCI \in QCI_{s_{GBR}} \\ 0.1 & \forall QCI \in QCI_{s_{Non-GBR}} \end{cases} \\
 Budget(QCI, T_{request}, T_{current}) &= \frac{T_{current} - T_{request}}{QCI_{linkBudget}} \\
 QoS(QCI) &= \left\{ QCI \in (QCI_{s_n}) / \frac{n+1}{|(QCI_{s_n})|+1} \right\} \\
 SpectralEfficiency(SelectedSlot, AvailableSlotCQIs) &= \\
 \left\{ = \frac{k+1}{|AvailableSlotCQIs|+1} \quad \exists k \in N, SelectedSlot_{CQI} \in AvailableSlotCQIs / SelectedSlot_{CQI} = AvailableSlotCQIs_k \right. \\
 Justice(NumSlotsAllocated, NumSlotsAllocatedUE, NumAvailableSlots) &= \\
 &= \frac{NumSlotsAllocated - NumSlotsAllocatedUE}{NumAvailableSlots} \\
 Reward &= \begin{cases} -1 & SelectedSlot \notin AvailableSlots \\ GBR \times Budget \times QoS \times SpectralEfficiency \times Justice & SelectedSlot \in AvailableSlots \end{cases}
 \end{aligned}$$


This system has perfect value if all characteristics are 1, anything out, i.e. a miss, will make the reward go down in the direction where it is failing, another good thing is that weights can be put in the actual Reward equation to give priority to some attributes.

The current state of this work is at verifying different traffic profiles that should provide datasets of good value for the context of 5G-Range project. A study conducted in South Africa has shown traffic characterization in that country, this work analyzed traffic for 10 days and has found that 68.45% is Web Traffic and the next big portion is Video, such as Youtube (Buffered) or VoIP systems (livestreams) [5]. Based on that study a dataset is being created covering the percentages of internet usage and also the user per kilometer requirement of the 5G Range Project of 2 users/km.

10

References

- [1] Okvist, Simonsson and Asplund. *LTE Frequency Selective Scheduling Performance and Improvements Assessed by Measurements*. International Symposium on Personal, Indoor and Mobile Radio Communications, 2011.
- [2] Yang, Xu, Han, Rehman and Tao. *GA Based Optimal Resource Allocation for Device to Device*. WCNC 2014 - Workshop on D2D and Public Safety Communications, 2014.
- [3] 3GPP. *Policy and Charging Control Architecture*. TS 23.203, 2018.
- [4] Arulkumar, Deisenroth, Brundage and Bharath. *A Brief Survey of Deep Reinforcement Learning*. IEEE Signal Processing Magazine Special Issue on Deep Learning for Image Understanding, 2017.
- [5] Johnson, Pejovic, Belding and Stam. *Traffic Characterization and Internet Usage in Rural Africa*. WWW'11 Conference companion on World Wide Web, 2011



Competitive European/Brazilian SMEs

Cloud-based time-limited transaction management

Author: Roberto C. Mayer (MBI and BraFip)

Co-author: Antonio Moreno (MBI)

Who stands to benefit and how

Software developers and researchers building new cloud-based software solutions based on transaction processing.

Position Paper

Cloud Computing and Internet of Things have opened up a new space for innovation opportunities. However, new challenges present themselves to solution developers.

Not only does every user of a mobile device have access to as many cloud-based services as possible, but the rapidly growing number of CPUs deployed on our planet soon will surpass the number of human fingers, hence requiring machine-to-machine communication, integrating all of these with what is called 'legacy IT'.

Additionally, the cloud computing environment imposes additional challenges to every non-trivial solution to be deployed massively:

1. Application Servers need to be protected from users accessing them from unstable connections, as well as from end-user devices that simply go off due to dead batteries.
2. Need to be resilient against Distributed Denial of Services attacks.
3. Sharing of user experience data among diverse interactions channels (web, mobile or physical installations).
4. Be able charging users at very small amounts per transaction (micropayments).
5. Allow cloud-based access to corporate servers without risking low availability to internal users.
6. Create different priorities for IoT data traveling across the cloud, according to being critical or not.
7. Managing many similar versions of solution running simultaneously.
8. Auditability & proof of compliance with SLAs.
9. Running same code in multiple languages & cultures.
10. Provide developers with significant information when things don't run as expected ("bugs").

Studying these, we concluded that it is economical nonsense to require each solution development project to pay for the costs of solving these problems again and again, as they are requirements imposed by the common environment. Hence, we designed and implemented a cloud-based transaction manager, called VersaCloud, to provide solutions to the above challenges as cloud-based services.

In short, VersaCloud combines features of API managers and online transaction processing engines, plus a bunch of additional features, into cloud-based middleware.

VersaCloud introduces two modifications to the traditional database transaction concept.

First, transactions are composed of individual operations to be provided by as many cloud-based servers as necessary (instead of various database tables). Secondly, all transactions are time-limited: if not committed in a previously specified period of time, they are automatically rolled back (and backend servers are notified, if necessary).

VersaCloud is free to use by developers and researchers. Its technical documentation is available at <https://www.versacloud.technology>.

VersaCloud's core ideas have been object of an US Patent request.

Support SMEs in becoming competitive and exploiting the potential of international markets: EU-Brazil SME coo

Author: Jacques Magen (Interinov, Chair of the 5G PPP SME Working Group)

Position Paper

Small and Medium-sized Enterprises (SMEs) have an important role to play in developing, piloting and deploying 5G technologies, both to help with disruptive technologies and to address the needs of various vertical sectors.

SMEs in Europe are bringing a great added value in providing innovative concepts and solutions that are having an important impact on the 5G value chain. They have the agility and flexibility required in a fast evolving technical and market landscape. This has been well understood in the 5G Public Private Partnership (5G PPP), where SMEs have an a 20% share.

5GINFIRE is a 3-year research and innovation action with partners from Brazil: University of São Paulo and Federal University of Uberlândia contributing towards an open, and extensible 5G NFV-based Reference (Open5G-NFV) ecosystem of Experimental Facilities, supporting the digitisation of vertical industries. The 5GINFIRE platform for innovation is specifically suitable for SMEs.

European SMEs working in the domain of 5G have been benefiting since 2015 of the launch of the 5G Infrastructure Public-Private Partnership, in short 5G PPP (<https://5g-ppp.eu/>). The 5G PPP is a joint initiative between the European Commission and European ICT industry (ICT manufacturers, telecommunications operators, service providers, SMEs and research Institutions). The 5G PPP aims to deliver solutions, architectures, technologies and standards for the ubiquitous next generation communication infrastructures of the coming decade. The 5G PPP will open a platform that helps reach the common goal to maintain and strengthen the global technological lead of Europe. A total investment of 700 M€ is planned between 2014 and 2020.

The first phase of the 5G PPP focused on researching technology for 5G. The second phase, launched in 2017, is looking at ensuring the feasibility of using the 5G technology and solutions in various vertical sectors such as content & media, energy, transport & logistics, health and others. The third phase, which is being kicked off in the second half of 2018, will concentrate on providing a set of 5G infrastructures that will be used for actual large-scale trials in various vertical sectors. A set of “Key Performance Indicators” has been set for the 5G PPP. One of those indicators is the participation of SMEs in the 5G PPP, which should exceed 20%. In Phase 1 and Phase 2, SME participation represents about 19% of the budget, almost reaching the objective.

A major effort has been made to promote SME participation in the 5G PPP and to make the expertise of SMEs more visible, under the coordination of a dedicated “SME Working Group”. This Working Group is open to all SMEs working in 5G and related domains, via the NetWorld2020 European Technology Platform (<https://www.networld2020.eu/>). Activities related to the Working Group included the coordination and organisation of SME participation in major events (e.g. via booths and presentations), the production of specific SME brochures, and a section dedicated to promoting the skills and expertise of the SMEs on the 5G PPP and NetWorld2020

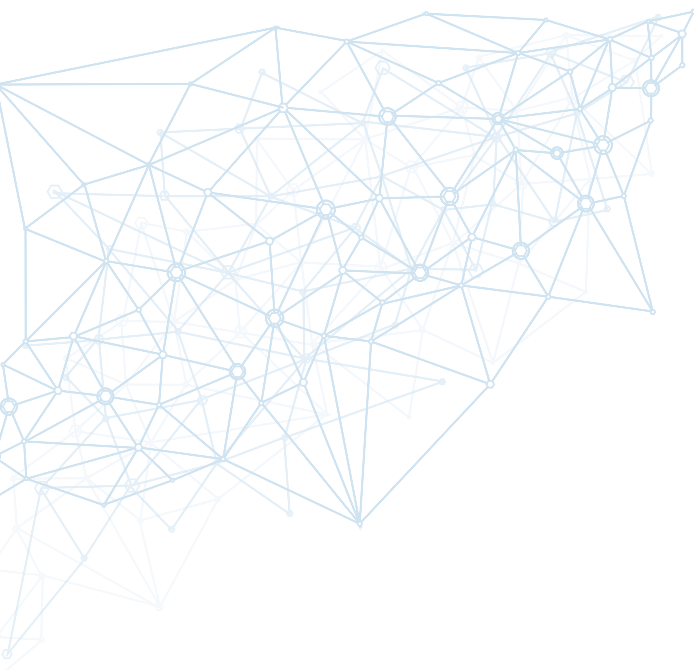
web sites (<https://www.network2020.eu/sme-support/>). Announcements and promotion has been made regularly to the broader community via various communication channels.

Considering the current evolution towards 5G trials in Europe and all over the world, the involvement of SMEs, to support the deployment of 5 infrastructures and to test 5G related solutions in various vertical sectors, is more than ever critical. In this context, SME participation in the 5G PPP is expected to grow.

5G deployment and trials need to happen at global scale. Therefore, the 5G PPP has created a dedicated “Trials Working Group”, which in turn includes several sub-groups. One of them is focusing on international 5G trials.

The 5G Infrastructure Association, the voice of the European industry in 5G, who represents the private side in the 5G PPP, is part of a multilateral Memorandum of Understanding that involves several organisations all over the world, including 5G Brazil. The next “Global 5G event”, co-organised by all organisations involved, will be hosted in Rio de Janeiro on 28-30 November 2018. 5G Brazil has declared their intent to participate in global trials (<https://www.sdxcentral.com/articles/news/5g-brazil-project-focuses-on-trials-global-collaboration/2018/04/>). The European Union and Brazil have also signed an agreement to strengthen their cooperation in 5G (<https://5g-ppp.eu/eu-and-brazil-to-work-together-on-5g-mobile-technology/>). In this context, strengthening the collaboration between Europe and Brazil at the time when 5G infrastructures are being deployed and actual 5G trials are starting should be one of our objectives.

We propose here to work on initiating a cooperation allowing European and Brazilian SMEs to work together and with larger organisations, both on innovative 5G technologies to be implemented in 5G infrastructure platforms, and on trials related to 5G solutions for various vertical sectors. This should happen both in Europe and in Brazil. The mechanism to implement this cooperation could be inspired on the one hand by the work performed by the SME Working Group in Europe, and on the other hand by the international cooperation already in place between the 5G Infrastructure Association and 5G Brazil.



16

Digitizing Industry



The Impact of 5G on Vertical Industry Sectors

Author: John Favaro (Trust-IT Services and Global5G.org)

Co-authors: Stephanie Parker (Trust-IT Services and Global5G.org), Andrea Schillacci (Trust-IT Services and Global5G.org)

Who stands to benefit and how

All those involved in a number of key industrial sectors, especially automotive, agribusiness, and smart cities, who need to understand the impact that the fifth generation (5G) wireless technology will have on the digitisation of their sector.

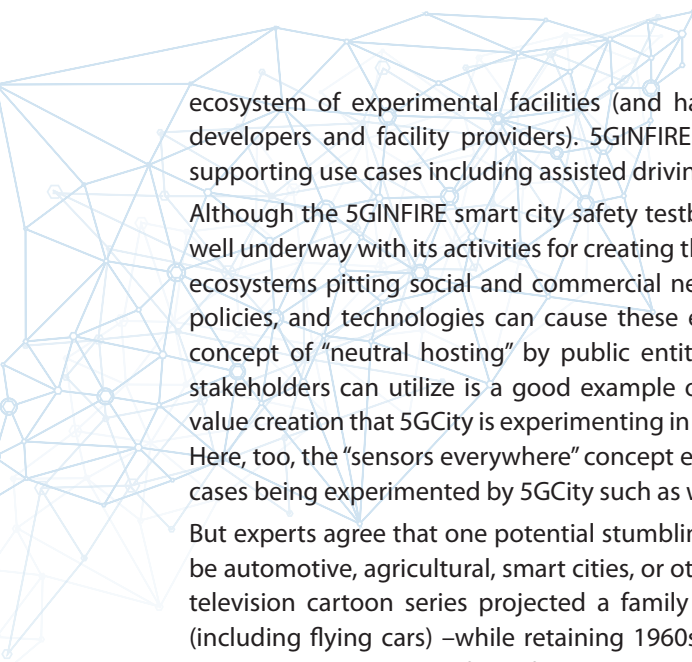
Position Paper

Fifth Generation (5G) wireless network technologies are destined to play a transformative role in our daily lives, bringing us myriad new services on our smartphones, tablets, and televisions. But the truly transformative effect of 5G will happen through its role in digitizing industry. The ultra-fast, reliable connectivity of 5G is an essential enabler of Industry 4.0, providing the capacity to handle the Big Data and Cloud computing that is at the core of its vision. Its wireless nature will enable cost-effective connection of thousands of devices, changing the way that (for example) sensors are utilized in factories. The concept of “sensors everywhere” is no longer a fantasy, and will lead to a complete rethinking of manufacturing architectures and processes.

5G-enabled “sensors everywhere” will also revolutionize the automotive vertical sector. The modern car’s onboard sensors will soon be flanked by hundreds, even thousands of connected sensors not only in the vehicle itself but all along “corridors” of 5G-equipped road infrastructure. These sensors will utilize 5G’s high-bandwidth, low-latency communication features to enable a quantum leap in “situational awareness” both for human drivers and self-driving cars: vehicles will be able to warn each other about their own actions (e.g. warning the car behind you that an Automatic Emergency Braking action has just been initiated), and the road infrastructure will be able to warn vehicles that a vulnerable pedestrian has suddenly come into range. But at the recent conference EUCNC 2018 in Ljubljana, the 5GCAR project noted that such a collaboration can only occur if all stakeholders are involved, from automobile manufacturers through infrastructure operators and technical equipment suppliers. In their recent (February 2018) White Paper, the 5G Automotive Working Group clarified this complex web of relationships among the various stakeholders and analyzed their potential contributions to successful 5G deployment for automotive use cases, both in societal and economic terms.

Many will be surprised to learn, however, that the transition to fully automated driving will most likely not occur in the automotive sector at all, but rather in agribusiness (think “simpler traffic context”). In fact, agribusiness may well stand to profit even more from 5G than the automotive business, with full connectivity opening the door to Big Data analytics for crop planning, energy savings, monitoring of soil and weather conditions, and improved predictive maintenance. Agricultural vertical use cases have top priority in European and Brazilian strategic planning.

There is an urgent need for a uniform testbed / platform to provide common technologies, share expertise, and leverage experience as it is gathered by the verticals. Maximum use of industry-leading and open source technologies helps to contain potentially prohibitive experimentation costs. 5GINFIRE, with both European and Brazilian partners, is building an open and extensible



ecosystem of experimental facilities (and has an ongoing open call for more open source developers and facility providers). 5GINFIRE includes an IT-AV Automotive Environment for supporting use cases including assisted driving using Virtual Network Functions.

Although the 5GINFIRE smart city safety testbed is still in its early stages, the 5GCity project is well underway with its activities for creating the hyperconnected city. Cities give rise to delicate ecosystems pitting social and commercial needs against each other, and mistaken decisions, policies, and technologies can cause these ecosystems to wilt as quickly as they arose. The concept of “neutral hosting” by public entities of critical 5G infrastructure that commercial stakeholders can utilize is a good example of the balancing act between public and private value creation that 5GCity is experimenting in the municipalities of Barcelona, Bristol, and Lucca. Here, too, the “sensors everywhere” concept enabled by 5G is making possible entirely new use cases being experimented by 5GCity such as waste dumping prevention in urban areas.

But experts agree that one potential stumbling block for all 5G vertical sectors – whether they be automotive, agricultural, smart cities, or others – is the so-called “Jetson effect”. This popular television cartoon series projected a family into the future with next-century technologies (including flying cars) –while retaining 1960s societal and cultural norms (such as hazardous woman drivers and the fact of even having human drivers at all). To avoid the Jetson effect, vertical industries must envision a future not only of 5G-enabled technological innovation, but of innovation at 360 degrees, including new business and payment models, new forms of ownership, and new types of public / private collaboration.

Related Demos at Cloudscape Brazil 2018: FarmCloud, SME agribusiness, <http://www.farmcloud.io/> | <https://www.atmosphere-eubrazil.eu/multiple-farm-controllers-single-cloud-platform>

FASTEN – Flexible and Autonomous Manufacturing Systems for Custom-Designed Products

Author: Ariane Rodrigues Pereira (INESC Brasil)

Co-author: Gustavo Dalmarco (PUCRS), Symone Gomes Soares Alcalá (UFG)

Position Paper

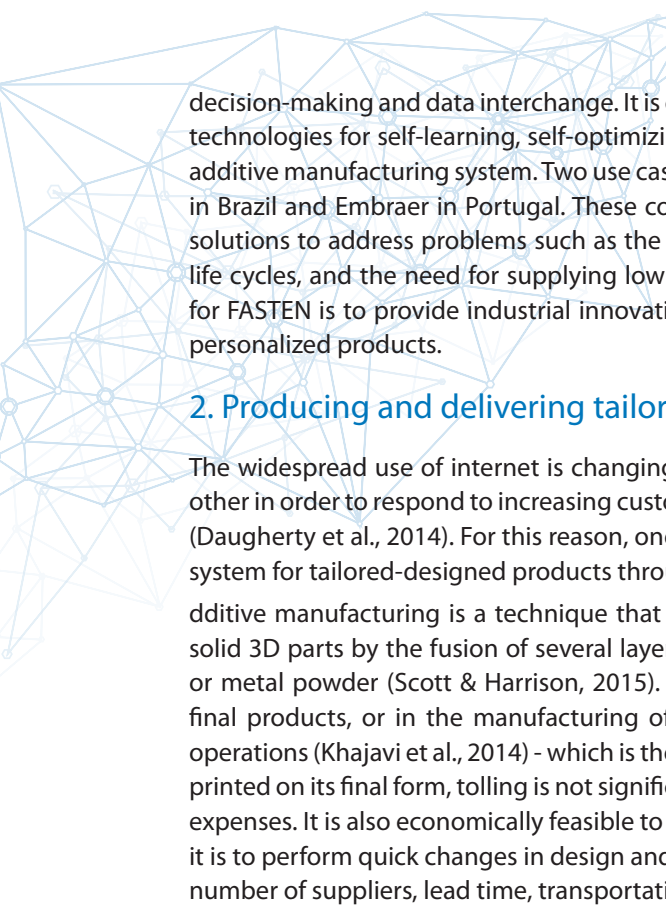
The use of IoT on different industries is establishing a new technological paradigm for manufacturing and sales. The possibilities of the new Industry 4.0, fostered by additive manufacturing, goes from a web-based integration of production lines to the possibility of high quality personalized products. In the Brazilian scenario, however, the use of such technologies is moving slowly. In this sense, the promotion of IoT solutions through joint university-industry activities may be the main path towards disseminating the possibilities of this technology, including flexible manufacturing through a robotic additive manufacturing approach.

According to the Brazilian National Confederation of Industry (CNI), the intensive use of digital technologies in Brazilian industry is less widespread. In 2016, about 58% of the overall amount of industries were aware of this importance for competitiveness, but less than half use these technologies. In addition, this maximum organization of the Brazilian industrial sector states that Industry 4.0 can be leveraged in Brazil, for example, through the creation of demonstration platforms that encourage and motivate companies to adopt digital technologies for industrial production. Companies which invest and become experts in the Industry 4.0 fundamentals certainly will become more competitive. Initiatives on these challenges can show to the industries that Advanced Manufacturing, the Industrial Internet of Things and Additive Manufacturing and Robotics, among others innovative areas, allow to integrate physical and digital technologies, to combine product development, manufacturing and logistics planning, and to link systems, machines and human labour.

The referred innovative solutions can bring positive effects on efficiency and quality of products and services fostering an improvement of productivity, flexibility, and profitability, among others business demands for reaching competitiveness. Therefore, initiatives such as FASTEN can promote the conditions to consolidate a breeding environment for Industry 4.0 that allows to strengthen the Brazilian industry's sustainability.

1. Introduction

The advent of the Industry 4.0 is posing several challenges for industries - for instance the pace of change, technologies to adopt, and user integration into the development process. In this context, digital technologies such as cloud infrastructures, big data and artificial intelligence, along with physical advancements in smart materials, nanotechnology and 3D printing play a key role in such fast developing scenario (Kumar, 2018). In line with the Industry 4.0 new paradigm, FASTEN project aims to develop, demonstrate, validate and disseminate an integrated and modular framework for efficiently producing custom-designed products. Hence, FASTEN will demonstrate an open and standardized framework to produce and deliver tailored-designed products, capable to run autonomously and deliver fast and low-cost additive manufactured products. This is expected to be achieved by effectively pairing digital integrated service/products to additive manufacturing processes, on top of tools for decentralizing



decision-making and data interchange. It is considered the application of sophisticated software technologies for self-learning, self-optimizing, and advanced in order to build a full connected additive manufacturing system. Two use cases benefit of FASTEN project, they are ThyssenKrupp in Brazil and Embraer in Portugal. These companies will have the opportunity to have flexible solutions to address problems such as the increasing demand diversity, products with shorter life cycles, and the need for supplying low volumes per order. The overall benefits envisioned for FASTEN is to provide industrial innovations capable to effectively manufacture and deliver personalized products.

2. Producing and delivering tailored-designed products

The widespread use of internet is changing the way supply chain echelons interact with each other in order to respond to increasing customer requests of personalized products and services (Daugherty et al., 2014). For this reason, one of FASTEN activities is to develop a manufacturing system for tailored-designed products through Additive Manufacturing (AM).

Additive manufacturing is a technique that consists of the reproduction of CAD drawings into solid 3D parts by the fusion of several layers of a specific material, either in the form of plastic or metal powder (Scott & Harrison, 2015). The 3D parts can be used as parts of an assembly, final products, or in the manufacturing of parts for Maintenance, Repair and Overall (MRO) operations (Khajavi et al., 2014) - which is the application foreseen by FASTEN. Since AM parts are printed on its final form, tolling is not significantly necessary, thus reducing production time and expenses. It is also economically feasible to produce small batches (especially one-of-a-kind), as it is to perform quick changes in design and customization. AM has also potential to reduce the number of suppliers, lead time, transportation services and inventories (Holmström et al, 2010).

20 3. Delivering fast and low-cost additive manufactured products

To cope with an increasing demand diversity, products with shorter life cycles, and low volumes per order, manufacturing companies need flexible solutions, capable to effectively manufacture and fast deliver low-cost personalized products. With the FASTEN application, the objective is to accelerate the integration between products design and manufacturing, effectively controlling and assessing the performance of the whole manufacturing systems. This integration will be achieved by combining simulation, optimization, and analytics tools, to virtualize the whole system, thus providing a high degree of flexibility and agile decision-making in real-time. These technologies will effectively approximate customers to products, and significantly reduce the global operational costs, leading to low unitary costs of custom-designed products and fast deliveries.

4. Concluding statement

FASTEN can have an enormous impact on the Brazilian and Portuguese industrial economies and markets by introducing a framework for dealing with Industry 4.0. It will use a standardized framework based on available standards - hardware and open-source software - customized to deliver fast and low cost additive manufactured parts for efficiently producing tailored-designed products.

5. References

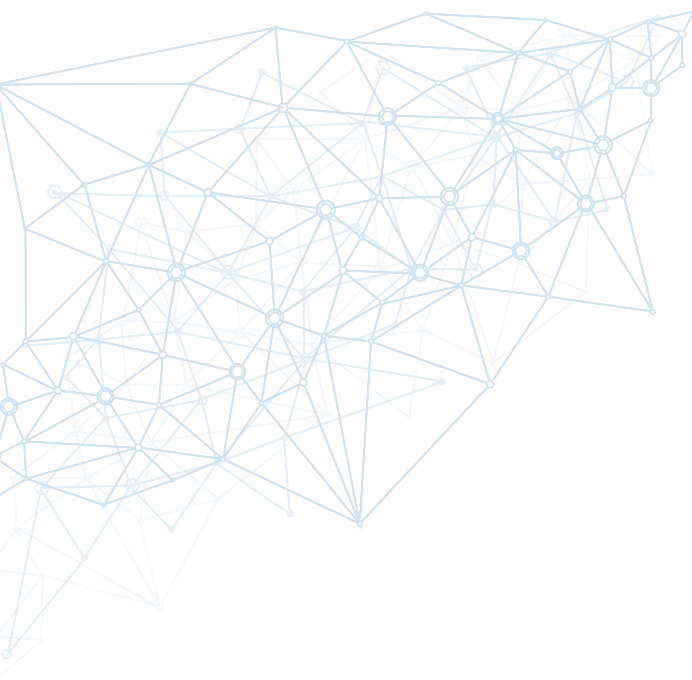
Daugherty, P., Banerjee, P., Negm, W., & Alter, A.E. (2014) Driving Unconventional Growth Through the Industrial Internet of Things. Accenture Technology. Retrieved from <http://www.mcrockcapital.com/uploads/1/0/9/6/10961847/accenture-driving-unconventional-growth-through-iiot.pdf>.

Holmström, J., Partanen, J., Tuomi, J., & Walter, M. (2010). Rapid Manufacturing in the Spare Parts Supply Chain: Alternative Approaches to Capacity Deployment. *Journal of Manufacturing Technology Management*, 21(6), 687-697.

Khajavi, S. H., Partanen, J., & Holmström, J. (2014). Additive Manufacturing in the Spare Parts Supply Chain. *Computers in Industry*, 65(1), 50-63.

Kumar, A. (2018) Methods and Materials for Smart Manufacturing: Additive Manufacturing, Internet of Things, Flexible Sensors and Soft Robotics. *Manufacturing Letters* 15, 122-125.

Scott, A., & Harrison, T. P. (2015). Additive Manufacturing in an End-to-End Supply Chain Setting. *3D Printing and Additive Manufacturing*, 2(2), 65-77.



EU-Brazil Common standards

A Proposal to Apply a Risk Assessment Methodology for IoT Systems to a Smart Childhood Obesity Caring Solution

Author: Sergio Luis Ribeiro (CPqD)

Co-authors: Emilio Nakamura (CPqD), Rodrigo Lima Verde Leal (CPqD)

Position Paper

This paper presents a risk assessment methodology focused on privacy, security, safety, resilience and reliability to IoT systems. The methodology is composed of ten phases that comprehends the risks elements to calculate the probability of a threat agent to explore one or more vulnerabilities in an IoT asset that turns a threat into an incident that causes impacts on different actors: manufacturers, developers, customers, integrators, service providers and users.

Introduction

IoT systems require different security levels depending on their specific use case. Attacks in any of their assets may cause global impacts. The proposed risk assessment methodology takes the use case point of view for the IoT system to provide an integrated security view that directs the actions to be taken by the different actors. For instance, sensors or actuators manufacturers and IoT platform or applications developers can implement the necessary security controls highlighted by that specific use case. Integrators and service providers can build an IoT system using only the assets that comply with the necessary security levels. In addition, users can choose the best IoT system based on the provided security requirements. IoT systems have key objectives regarding trustworthiness [1]: privacy, security, safety, reliability and resilience. The proposed methodology focuses on these objectives to calculate an IoT system risks. It is desirable that this methodology is applied to a smart childhood obesity caring solution being developed in project OCARIoT.

23

Risk Assessment Methodology

Proper risk assessment considers the probability of a threat agent exploring a vulnerability in an asset, turning a threat into an incident that represents an impact. The starting point of the methodology is to consider that it is not possible to protect against unknown or mistakenly assessed risks. As a part of risk management processes (ISO 31000:2018 [2]), risk assessment provides guidance to define and implement security controls that are both efficient and effective. The limited available resources are used to treat the most important risks in an organized and formal way, without politics, personal preferences or interferences. Risk assessment can be performed in different contexts, according to the desired risk view and the involved actors. In IoT, there are views for sensors and actuators manufacturers, for platform, application and developers, for customers (home, health, city, industry), or for integrators, service providers and for users. The proposed methodology [3] has 10 phases (Figure 1) and defines the use case as the context for the risk assessment for all the actors. The main reason is the multiplied cyber-physical attack points resulting from integration between sensors, actuators, platforms, applications and users where an attack in one point affects the whole system. The

other reason is that a same asset can be used in different IoT contexts that require different security levels based on the specific risks involved.

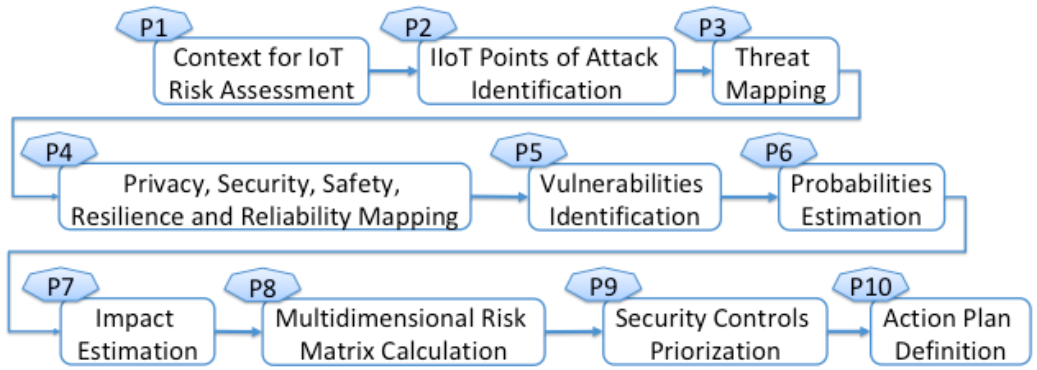


Figure 1: 10-phase methodology.

Project OCARIoT

The main objective of OCARIoT is to provide an IoT-based personalised coaching solution guiding children to adopt healthy eating and physical activity behaviour. The IoT network will allow us to observe child activity patterns of daily living, health evolution, physiological & behavioural parameters and environmental data. All this information combined with medical patterns will allow us to provide a customised obesity coaching plan while enabling children to remain active and engaged in their well-being and healthy habits management. OCARIoT will demonstrate and validate its results on three specific pilot sites in Spain, Greece and Brazil.

24

Preliminary Results

The project is still in its early stages. Since OCARIoT can generate, manipulate and store personal information, it is desirable, under the project's activities, to carry out a risk assessment throughout the Pilot's environment, in order to determine whether the proposed controls and best practices have been followed. Each of the abovementioned phases will be undertaken, considering the involved actors (children, educators and families) and scenarios (@school, @home and @city).

Acknowledgement

The authors acknowledge the financial support given to this work, under project OCARIoT, which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731533 and the RNP under No 3007. This paper reflects only the author's views and the Agencies are not responsible for any use that may be made of the information contained therein.

References

- [1] Industrial Internet Consortium, "Industrial Internet of Things Volume G4: Security Framework", IIC:PUB:G4:V1.0:PB:20160926, September, 2016, http://www.iiconsortium.org/pdf/IIC_PUB_G4_V1.00_PB-3.pdf. Accessed in Feb 1, 2018.
- [2] International Organization for Standardization, "ISO 31000:2018, Risk management – Guidelines", 2018.
- [3] Nakamura, Emilio; Ribeiro, Sergio. "A Privacy, Security, Safety, Resilience and Reliability Focused Risk Assessment Methodology for IIoT Systems". Global IoT Summit Proceedings. 2018.

Protecting sensitive data in cloud environments

Authors: Andre Carvalho (UFAM), Altigran Silva (UFAM), Ronny Guimarães (UFAM)

Co-authors: Christof Fetzer (TU Dresden) and Tanara Lauschner (UFAM)

Who stands to benefit and how

Cloud users, providers and Database Systems researchers.

Position Paper

In this age of ever-growing utilization of cloud resources for computation, the need for trustworthiness insurance from data management systems is even more important. Datasets grow ever larger, minimizing the possibility of processing data locally, especially in resource intensive tasks, such as data mining. Moreso, machine learning techniques such as Deep Learning usually entails manipulating large datasets.

In this context, and with the implementation of GDPR, there is a well deserved and urgent need for the protection of sensitive data. Sensitive data is any data that must be protected from unauthorized access to safeguard the privacy or security. This concept can be applied to individual data, such as medical records, social security numbers, biometric data, etc, and even to business-related information.

Protecting sensitive data is an essential task in data management. This task is even more important in cloud-based environments, where there is no external control over who has physical access to the cloud infrastructure. This indicates that the cloud provider could be a very serious attack vector.

For instance, a malicious attacker with physical access to the cloud can potentially read sensitive data, not only by reading data files in discs (which can be protected by being encrypted), but also by memory scanning. This means that at any given moment where the raw sensible data is loaded into main memory in a cloud machine it is vulnerable to attacks.

Moreso, even a non malicious user can inadvertently expose private data by performing queries that do not take into account the privacy of the individuals depicted in the results, releasing sensitive information to the public. An example of such exposure was the AOL search data leak, where then-believed anonymized data was released to the public, however many queries contained personally identifiable information, which lead to the identification of search queries from a number of real users. That scenario can be even more damning if the leaked data is something like, for instance, medical records.

Data protection in this context can be seen as having two axis: Data access and privacy. Data access is ensuring that no unauthorized party may access sensible data in any form, while privacy regards avoiding that personally identifiable information is present in results from queries on the sensible data, by properly anonymizing it.

The main challenge when dealing with sensitive data in the cloud is that it must not ever be readable by a third party in the cloud system. This mean that, at no moment, this sensitive data may be present in its raw, readable version, until it arrives at the destination user. While this poses no problem for simple object retrieval systems, many systems require a more sophisticated processing of the data, where usually this processing leads to manipulation of

data in memory. Traditional SGBDs may have the option to store the data in an encrypted form, but the processing of queries is done with the raw data.

One solution for this problem found in the literature is the use of techniques to perform queries on encrypted data. The main (and obvious) advantage of these methods is that there is no need to decrypt the sensitive data to perform such queries. However, these techniques are mostly focused on simple operators and keyword matching, which may still not be enough for more advanced scenarios, such as most that would need the use of a Relational SGBD.

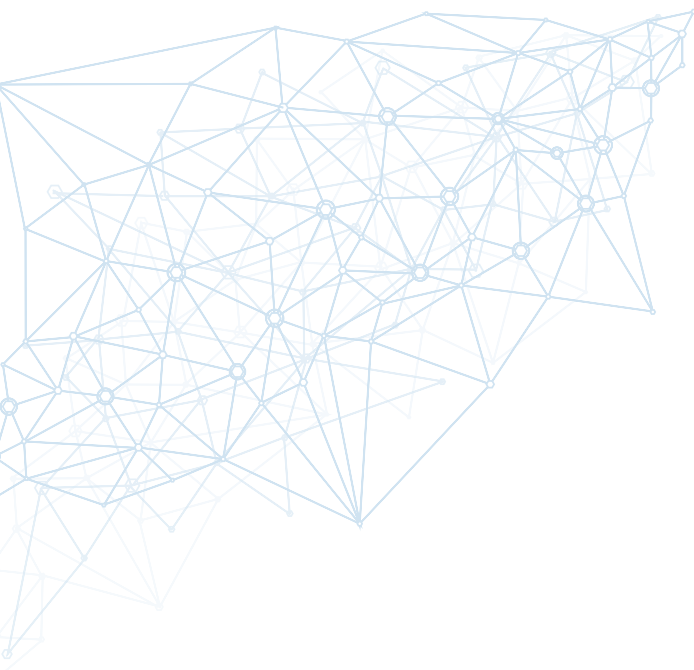
In 2015 Intel introduced Software Guard eXtensions (SGX) in their architecture. This enabled the use of SGX enclaves, runtime environments where the memory is encrypted and thus unreadable from outside attackers. We believe that the use of such enclaves might be a solution to avoid memory reading attacks in cloud servers.

In the scope of the ATMOSPHERE project (atmosphere-eubrazil.eu), we propose a solution for this problem: The creation of a secure data layer, whose objective is to assure that no sensitive data is at any point in its raw, readable state in the cloud infrastructure, the Atmosphere Data Layer (DaLay). Its objective is to intermediate all data access requests made by users or data processing services to some target sensitive database management service (TDBMS) that fits the application. The purpose of the DaLay is to isolate the data stored in the TDBMS from the clients accessing it, while also being responsible for the access control mechanisms for the data. Thus, it must grant access to data requests only to authorized parties, while also guaranteeing its privacy policies.

All data accesses are handled by DaLay, with the TDBMS running in isolation inside an enclave and all connections using end-to-end encryption with the key only known inside the enclave, making sure that the TDBMS does not accept any requests made in any different way. This will ensure that the data is not vulnerable to attacks stemming from having physical access to the cloud system and other man-in-the-middle approaches.

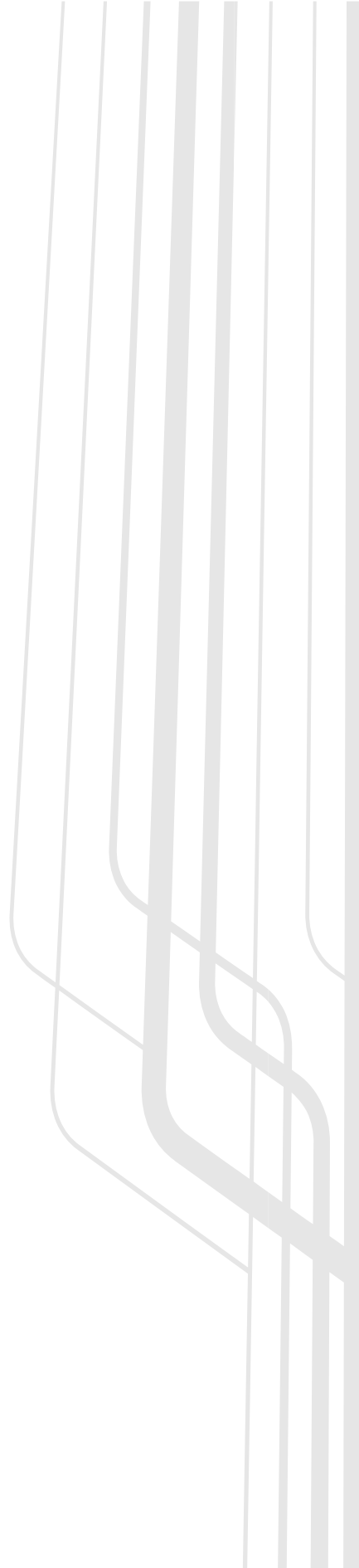
By its turn, the DaLay only sends to the TDBMS requests that satisfy the access and privacy policies that were defined by a curator. The policies are managed inside the enclave and stored in a encrypted form.

We believe that this approach has clear advantages: While it provides data encryption in all stages of query processing, it is agnostic to the data management system deployed. Moreover, it is agnostic to the data management system, and could theoretically be applied to most DBMSs. We are currently developing DaLay, and will soon release it to the community.



28

International Partnerships



What will the future hold for EU-BR collaboration in ICT?

Authors: Carlos Kamienski (Federal University of ABC), Juha-Pekka Soininen (VTT Technical Research Centre of Finland), Stenio Fernandes (Federal University of Pernambuco)

Who stands to benefit and how

Researchers: This position paper is expected to broaden the understanding researchers have of the nature of EU-BR coordinated calls in TIC. Also, it will encourage researchers to form consortia and submit proposals for future calls by providing some hints and tips they may help them in creating competitive proposals.

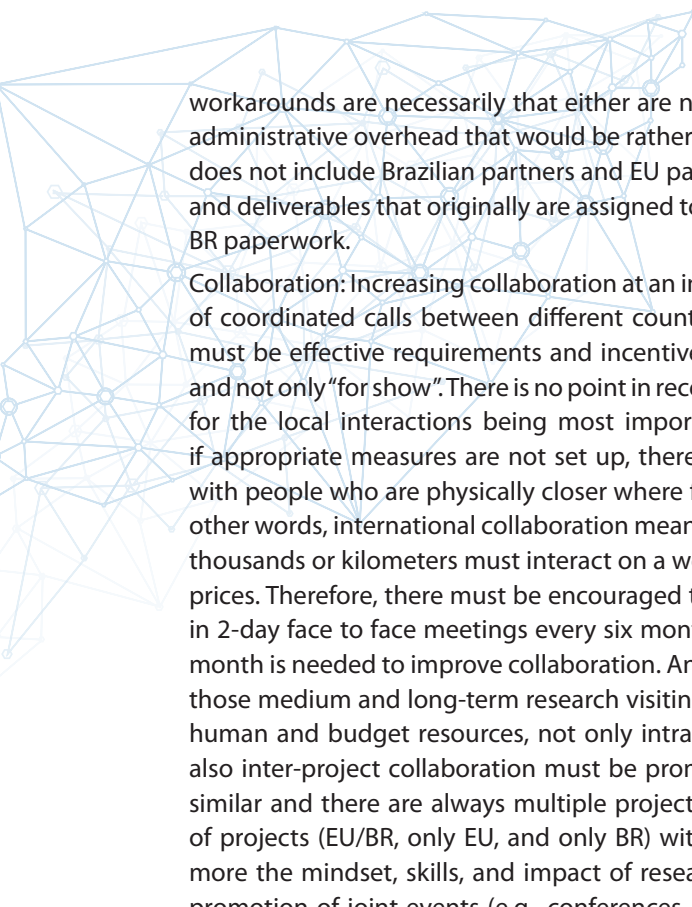
Policy Makers: Policy makers will have the opportunity to be exposed to the lessons learned by researchers who participated in different EU-BR projects, as well served as evaluators of project proposals and project reviewers.

SMEs: SMEs will have the opportunity of how they can benefit from taking part in consortia for future calls.

Position Paper

Context: Brazil and Europe have been launching coordinated calls for the development of collaborative research projects in the area of Information and Communication Technologies (ICT), focusing on topics such as Future Internet, Cloud Computing, Smart Cities, Big Data, 5G Networks and Internet of Things. Since 2010, four coordinated calls have been launched resulting in 193 proposals submitted involving hundreds of institutions (universities, research institutes, companies, government agencies) of both continents. From those proposals, 20 were selected where each party funded 25 million euros. In Brazil, these proposals are funded by the Ministry of Science, Technology, Innovation, and Communications (MCTIC) and run by RNP (the National Network of Research and Education), having the European Commission as its counterpart in Europe. The four calls were launched in 2010 (goo.gl/fpThxm), 2012 (goo.gl/XMTCfi), 2014 (goo.gl/ss4pmL) and 2016 (goo.gl/B6DVxK). Currently, projects selected in the first and second calls are already finished, projects selected in the third call are in progress and projects selected in the four calls started at the end of 2017. The outcomes in terms of research contributions and increased collaboration between Brazil and Europe have been positively assessed and therefore a fifth call is already being planned.

Challenges: However positive EU-BR research collaborations might have proven to be, there is always room to improve the legal and management frameworks in order to provide a better support for researchers to achieve their goals. Particularly, two key issues should be revisited as soon as possible. Firstly, even though the whole schedule is determined by an agreement between officers from both continents, one issue that always causes problems is the (lack of) synchronization of project start times due to the different legal frameworks in Brazil and Europe. For most Brazilian partners, the project officially starts and the financial resources are still not available, which is practice delays the real work to be done. Secondly, cooperation with other countries and regions requires additional flexibility in the existing legal and accounting frameworks. Both in Europe and in Brazil the documents, IT systems, and agreement templates were envisioned only for their local scope and when partners outside that scope participate,



workarounds are necessarily that either are not true with the original proposal or merely adds administrative overhead that would be rather unnecessary otherwise. For example, the EU DoA does not include Brazilian partners and EU partners must take responsibility for work packages and deliverables that originally are assigned to BR partners in the proposal. The same is valid for BR paperwork.

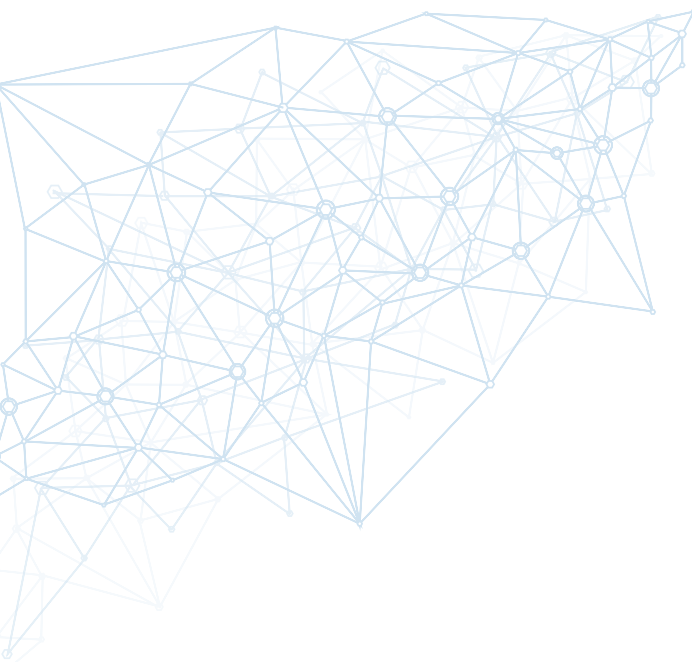
Collaboration: Increasing collaboration at an international level is the first and foremost purpose of coordinated calls between different countries and regions. At a policy-making level, there must be effective requirements and incentives to enforce international collaboration “for real” and not only “for show”. There is no point in receiving public funds for encouraging collaborations for the local interactions being most important than the international ones. Unfortunately, if appropriate measures are not set up, there is a human propensity to preferentially interact with people who are physically closer where face to face meetings happen more frequently. In other words, international collaboration means that partners who are separated by hundreds or thousands of kilometers must interact on a weekly or even daily basis. And distance charges its prices. Therefore, there must be encouraged that partners spend more time together, not only in 2-day face to face meetings every six month, but researcher exchange for periods over one month is needed to improve collaboration. An appropriate budget must be reserved to support those medium and long-term research visiting positions. Even considering the limited existing human and budget resources, not only intra-project collaboration must be encouraged, but also inter-project collaboration must be promoted. Since the relevant hot research topics are similar and there are always multiple projects selected for a particular topic, there are plenty of projects (EU/BR, only EU, and only BR) with which to collaborate in order to develop even more the mindset, skills, and impact of research activities. Given the inherent difficulties, the promotion of joint events (e.g., conferences, workshops, and exposition) is particularly a very effective approach for inter-project collaboration to flourish. Researchers from different projects and countries should have a playing field for sharing results, expertise and lessons learned and promoting the reuse of solutions from other projects.

Openness: The challenge of promoting and supporting collaboration, both at intra- and inter-project level, may be dealt with by firmly relying on open solutions. Here we propose the creation of a true open research ecosystem, based on multiple levels of openness, such as a) open source software; b) open data; c) open access; d) open innovation; e) open prototypes; f) open experiences; g) open experiments. Whereas some of these levels of openness are widely known, others are still in development. The key challenge here is how to promote such open research ecosystem where openness is a pervasive value and stimulates collaboration. Everything should be considered open by design, except for some specific cases such as privacy or intellectual property management.

Lessons: Since 2010 many people have had the opportunity to be exposed to and learn from international cooperation provided by the EU-BR coordinated calls in ICT, involving researchers, students, professionals, reviewers, and policymakers. The experience so far tells us that this is a very successful model that provided a good deal of lessons learned and insights gained, such as a) the effort made by the Brazilian government and the European Union for putting forward a public policy in a very strategic area that may contribute for the further development of Brazil and Europe is very important and should be publicized; b) cooperation between Brazil and Europe should be maintained and expanded; c) new players should appear and be encouraged to submit proposals to next EU-BR coordinated calls in ICT; d) this experience should be extended to other fields of knowledge; e) Europe has been working with coordinated calls for a while but at least for Brazilian policy makers, it is advisable to extend this style of joint calls

to other countries and continents (by the way, a coordinated call between Brazil and USA has already happened as a consequence of the successful cooperation with Europe); and f) potential barriers for new prospective researchers of both continents aiming to participate in new calls should be lowered and hints and tips should be provided on how to organize a consortium of Brazilian and European institutions.

Future: Future coordinated calls between Europe and Brazil should focus on the use of information and communication technologies for the big broad societal challenges of both continents, such as the impacts of climate changes, sustainability, circular economy, and urbanization. Specific applications or verticals to be addressed are water, energy, food, and air. ICT technologies to be used to deal with those challenges applied to particular verticals are IoT, Big Data Analytics, 5G, Artificial Intelligence, Quantum Computing, and Blockchain.



32

Regulation for Trust in Digital Environments

How much can I trust my cloud services?

Authors: Ignacio Blanquer (Universitat Politècnica de València) and Francisco Brasileiro (Universidade Federal de Campina Grande)

Co-authors: Danilo Ardagna (POLIMI), Andrey Brito (UFCEG), Amanda Calatrava (UPV), Andre Carvalho (UFAM), Vasiliki Diamantopoulou (UPRC), Christof Fetzer (TUD), Wagner Meira Jr. (UFMG), Regina Moraes (UNICAMP), Marco Vieira (UC)

Who stands to benefit and how

ATMOSPHERE aims at providing quantitative scores for such trustworthiness properties, so application developers and data scientists can build trust in their services based on pieces of evidence.

Position Paper

Trust is a choice that is based on past experience. Trust takes time to build, but trust can disappear in a second. Trusting cloud services is as complicated as trusting people. You need a way to measure it and pieces of evidence to build trust.

Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring, Hybrid Ecosystem for REsilient Cloud Computing (2017-2019) (hereinafter “ATMOSPHERE”) is a 24-month Research and Innovation Action, funded by the European Commission under the H2020 Programme and the Secretary of Politics of Informatics (SEPIN) of the Brazilian Ministry of Science, Technology, Innovation and Communication (MCTIC) that aims at designing and developing a framework and a platform to implement trustworthy cloud services on a federated intercontinental hybrid resource pool.

Trust in a cloud environment is considered as the reliance of a customer on a cloud service and, consequently, on its provider. Based on the given definition of trust in cloud computing, trustworthiness can be defined as the worthiness of a service and its provider for being trusted.

But trust lies on a broad spectrum of properties such as Security, Privacy, Coherence, Isolation, Stability, Fairness, Transparency and Dependability.

Nowadays, few approaches deal with the quantification of trust in cloud computing. ATMOSPHERE will support the development, build, deployment, measurement and adaptation of trustworthy cloud resources, data management and data processing services, demonstrated on a sensitive scenario of distributed telemedicine.

To achieve cloud computing trust services, ATMOSPHERE focuses on providing four components:

- » A dynamically reconfigurable federated infrastructure that provides isolation, high-availability, Quality of Service and flexibility for hybrid resources, including virtual machines and containers.
- » Trustworthy Distributed Data Management services that maximise privacy when accessing and processing sensitive data.
- » Trustworthy Distributed Data Processing services to build up and deploy adaptive applications for Data Analytics, providing high-level trustworthiness metrics for computing fairness and explainability properties.
- » An evaluation and monitoring framework, to compute trustworthiness measures from the metrics provided by the different layers, and trigger adaptation measures when needed.



The different trustworthiness properties identified need to be considered at different layers:

- » Security. Services should be resilient to malicious attacks and free of vulnerabilities. There are many vulnerability databases that can provide a score for applications, infrastructure and services. Security must be assessed at design time and reassured when new vulnerabilities come up, preventing access to vulnerable services and resources and migrating services to updated resources.
- » Privacy assurance. Privacy is the guarantee of an entity to be secure from unauthorized disclosure of sensible information. Privacy is a crucial property for trust, as privacy may not be guaranteed even in a secure environment with anonymised data. Computing the privacy risk, as the risk of reidentifying the subjects or inferring sensitive information, is needed to trigger more in-depth anonymisation techniques or to prevent the disclosure of the information.
- » Coherence. Distributed environments are highly convenient for high-availability, legal restrictions on data transfers and robustness, but consistency is a challenge. Assuming that eventual consistency is the feasible approach, one should know the degree of inaccuracy of such consistency.
- » Isolation. Multitenancy on a shared resource pool is a useful approach for enhancing resource utilisation and economy of scale. However, resource share inevitably implies isolation risks, especially in the performance of shared resources such as cache memory, network performance or disk I/O. Moreover, container-based computing reduces the isolation at the level of the process. Measuring how applications are mutually affecting themselves is very important to anticipate Quality of Service reductions and unavailability.
- » Fairness. Artificial Intelligence systems rely on the data used for the training and many complex parameters. Classification systems can be biased toward sex, race, education, etc. It is essential to understand this bias (which could be reasonable if the reality presents it) to interpret the results.
- » Transparency. Nowadays, we assume the results of Artificial Intelligence algorithms as ground truth, but in many cases, and especially in Deep Learning, results are hard to explain and understand. We do not know the critical parameters that drive our models, which could be vital in defining the liability of complex decision-making systems. Being able to understand those principles will enable defining the boundaries for a model.
- » Dependability. Differently from other trustworthiness properties discussed above, dependability has been better studied in the literature. It includes multiple sub-dimensions, such as Integrity, Availability, Reliability, Maintainability, Safety and Performance stability over time.

More information about ATMOSPHERE can be found in the website (<https://www.atmosphere-eubrazil.eu/>), in Twitter (@AtmosphereEUBR) and LinkedIn (<https://www.linkedin.com/in/atmosphere/>).

Fairness and Transparency in Trustworthy Cloud-based Analytics Services

Authors: Leandro Balby (UFMG) and Flavio Figueiredo (UFMG)

Co-authors: Nuno Antunes (UC), Vasiliki Diamantopoulou (UPRC), and Wagner Meira (UFMG)

Who stands to benefit and how

ATMOSPHERE aims at providing quantitative scores for trustworthiness properties, so application developers and data scientists can build trust in their services based on pieces of evidence.

Position Paper

1. Introduction

Machine learning (ML) is nowadays ubiquitous, providing mechanisms for supporting decision making in any data rich scenario. This rise in importance of ML raises societal concerns about the trustworthiness of systems that depend on it. The highest accuracy for large datasets is often achieved by complex models that humans struggle to interpret creating a trade-off between accuracy and interpretability, both of which affect trust in the system. In this context, the new general data protection regulation (GDPR) demands that organizations take the appropriate measures to protect individuals' data, and use it in a fair and transparent fashion.

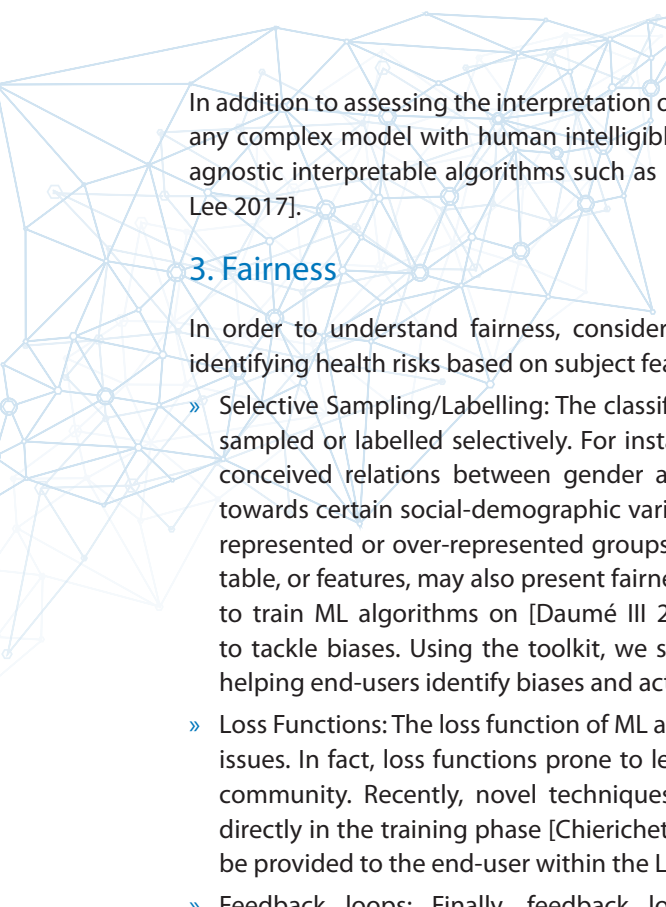
The ATMOSPHERE project (atmosphere-eubrazil.eu) considers trustworthiness as depending on many properties such as security, dependability, and privacy assurance among others. Moreover, data become a first class citizen, as trustworthiness also depends greatly on respecting data subject rights.

In the project, Fairness and Transparency emerge as key properties for the trustworthiness, with both terms being related. However, given the impact of ML systems on society, their definitions require care and may change according to the context. We view transparency as a means to capture interpretability of ML models. Complementary, fairness is related to both biases which may exist in datasets used to train ML systems, as well as biases which the ML system may incur on society.

2. Transparency

Interpretability can be defined as the degree to which a human can understand the cause of a decision [Miller 2017]. There is still no consensus on how interpretability should be assessed. In ATMOSPHERE, we adopt the view proposed by Doshi-Velez and Kim [Doshi-Velez and Kim 2017]:

- » Application level: The explanations are integrated into the outputs of the final model such that domain experts can evaluate it.
- » Human level: This corresponds to a simplified version of the application level. The difference is that here domain experts are not required, thus making experiments cheaper and more feasible.
- » Function level evaluation: Here humans are not required. This is more efficient when the models used are already well understood by humans.



In addition to assessing the interpretation of ML systems, we intend to promote it by equipping any complex model with human intelligible explanations. This is being done through model-agnostic interpretable algorithms such as LIME [Ribeiro et al. 2016] and SHAP [Lundberg and Lee 2017].

3. Fairness

In order to understand fairness, consider for instance a classifier trained with the task of identifying health risks based on subject features.

- » Selective Sampling/Labelling: The classifier above may be trained with data that was either sampled or labelled selectively. For instance, crowd-workers may label data based on pre-conceived relations between gender and health risks. Also, the dataset may be biased towards certain social-demographic variables. On the deployment phase, identifying under-represented or over-represented groups of subjects may mitigate such issue. Columns of a table, or features, may also present fairness biases. Some of such features can even be illegal to train ML algorithms on [Daumé III 2018]. Currently, we are using the Aequitas Toolkit to tackle biases. Using the toolkit, we shall evaluate the representation of subjects/labels, helping end-users identify biases and act accordingly.
- » Loss Functions: The loss function of ML algorithms may also be subject to biases and fairness issues. In fact, loss functions prone to less biases are gaining the attention of the scientific community. Recently, novel techniques exist which can deal with representation issues directly in the training phase [Chierichetti et al. 2017, Berk et al. 2017]. Such techniques can be provided to the end-user within the Lemonade framework.
- » Feedback loops: Finally, feedback loops occur when user actions guided by ML algorithms decrease trustworthiness. In our example, medical actions guided by an algorithm (e.g., a correlation between race and a certain risk), can increase biases when such data is used to re-train new models. Even though feedback loops rise due to user actions, it is possible to detect their presence [Ensign et al. 2017].

4. Concluding Remarks

Algorithms that measure transparency and fairness are currently being implemented using Lemonade [Santos et al. 2017]. Lemonade is a visual platform for distributed computing, aimed to enable machine learning applications. The platform will include a set of fairness or transparency metrics that are available to the user. We plan to evaluate fairness and transparency measurements on real world datasets from the medical domain. In particular, we are currently evaluating methodologies such as LIME, SHAP and Aequitas, already implemented on Lemonade, in both open and close domain data.

Device-Based Security to Improve User Privacy in the Internet of Things

Authors: Luis Pacheco (University of Brasilia), Eduardo Alchieri (University of Brasilia)

Co-authors: Priscila Solis (University of Brasilia)

Who stands to benefit and how

Cloud users and providers. Users are assured that their privacy requirements are met by the provider, while providers have means to implement such privacy requirements.

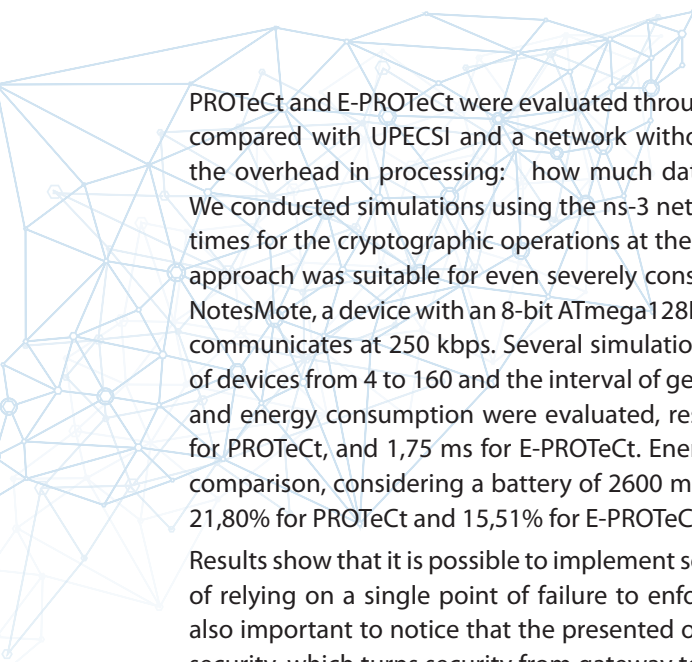
Position Paper

Internet of Things adoption is growing steadily and its integration with Cloud Computing is increasing the number of devices connected to the Internet. IoT devices, that are responsible for communication and actuation on user information, that are often sensitive, are exposed to several types of attacks, which may expose all other interconnected devices. Several proposals approach the security and privacy in the Internet of Things, most of them centralize the security protocols in a gateway device, which is responsible for aggregating the protection IoT data and sending it to the cloud. This approach brings several benefits, one very important is that it is possible to enforce security and privacy regardless of IoT networks technologies, which means that the protocol can be used in any situation.

In our work, we propose an architecture for privacy and the integration of the Internet of Things and Cloud Computing, adapting and moving the privacy and security mechanisms from the gateway to the IoT devices. This approach is aligned with efforts from several standardization organizations, that, in the last years, propose several communication standards for the Internet of Things with the goal of decreasing the fragmentation among current solutions.

Based in User-driven Privacy Enforcement for Cloud-based Services in the Internet of Things (UPECSI), the proposed architecture stores IoT data encrypted at the cloud platform, in this way only authorized entities (user and cloud services) have access to it, improving assurance of user privacy requirements. The approach of implementing security schemes at the IoT devices improves the architecture's fault-tolerance, since it removes a single point of failure. It also improves the overall security of the system, since there is no need of a component responsible for all user messages, which could impair the security properties of the system once compromised by a successful attack.

We called the proposed architecture as: PROTeCt: Privacy aRchitecture for integratiOn of internet of Things and Cloud computing. In our view, user's IoT networks will send data to the cloud to provide useful services to the user. When uploading data, IoT devices encrypt it with a symmetric key, which only authorized services have access. Keys are renewed periodically, enabling the addition and removal of services access to the data. Services are implemented through a Privacy Development Language (PDL), which requires to inform the actions taken with user data and enables users to enable and disable specific services features according to user's privacy requirements. A Trusted Third Party is responsible for auditing cloud services. We also propose Enhanced PROTeCt (E-PROTeCt), that decreases processing and transmission overhead by applying security schemes at the application layer instead of transport layer. This approach avoids encrypting data twice, since it must be encrypted for storage.

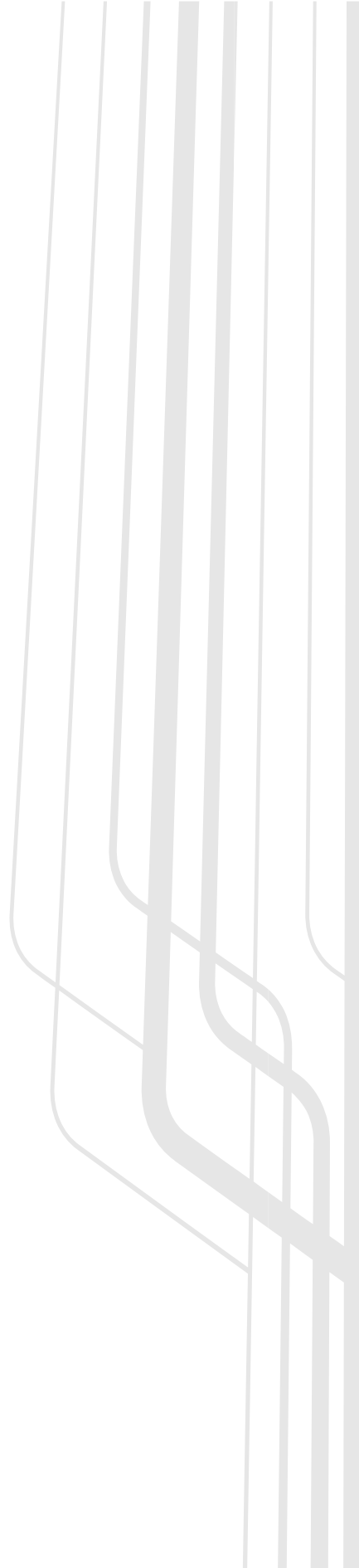
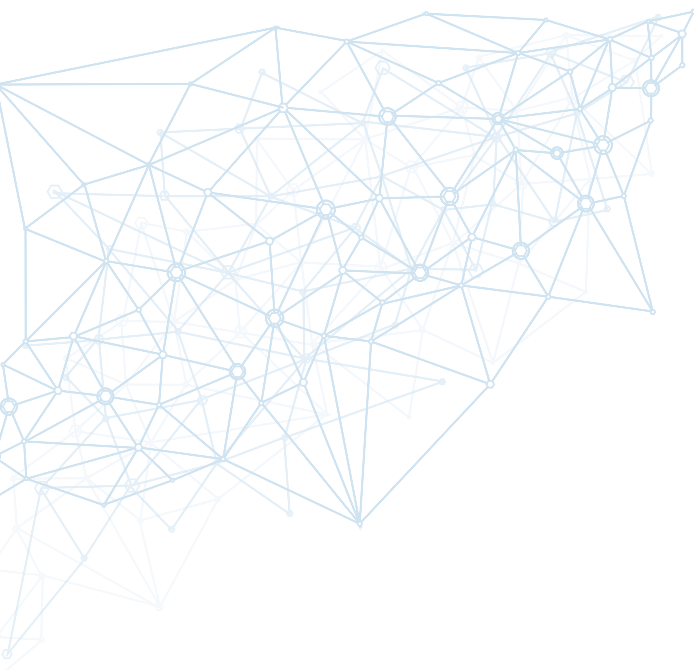


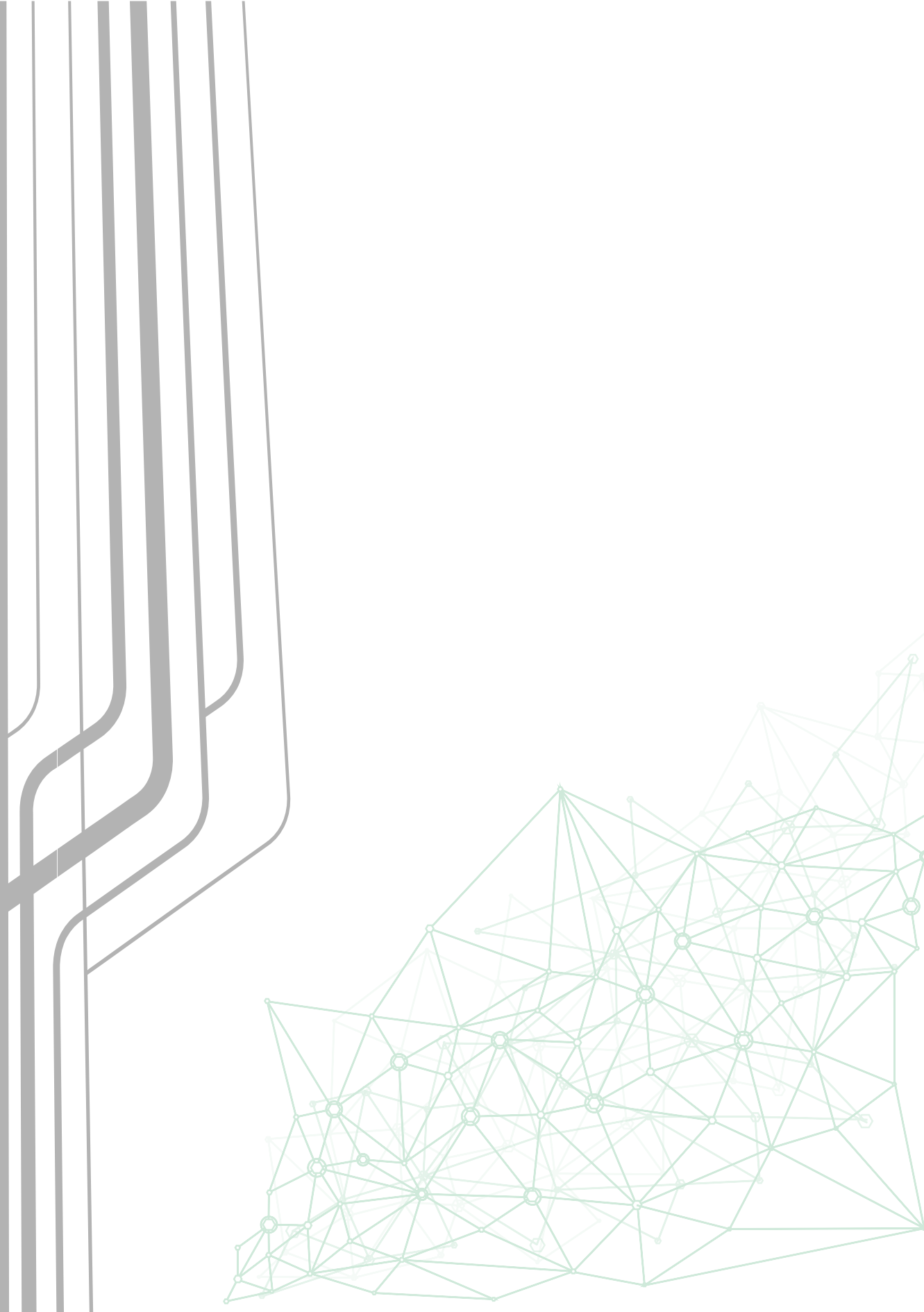
PROTeCt and E-PROTeCt were evaluated through analytical analysis and simulation, results were compared with UPECSI and a network without security. We derived equations to determine the overhead in processing: how much data was encrypted and how much data was sent. We conducted simulations using the ns-3 network simulator, we used experimentally acquired times for the cryptographic operations at the IoT devices and at the gateway. To show that our approach was suitable for even severely constrained devices, we simulated the use of a Micaz NotesMote, a device with an 8-bit ATmega128L microcontroller, 128KB RAM and 512KB ROM that communicates at 250 kbps. Several simulation scenarios were executed, varying the number of devices from 4 to 160 and the interval of generated data from 15 seconds to 5 minutes. Delay and energy consumption were evaluated, results show that delay overhead is around 3,5 ms for PROTeCt, and 1,75 ms for E-PROTeCt. Energy consumption was evaluated through lifetime comparison, considering a battery of 2600 mAh. In this case results show an overhead around 21,80% for PROTeCt and 15,51% for E-PROTeCt.

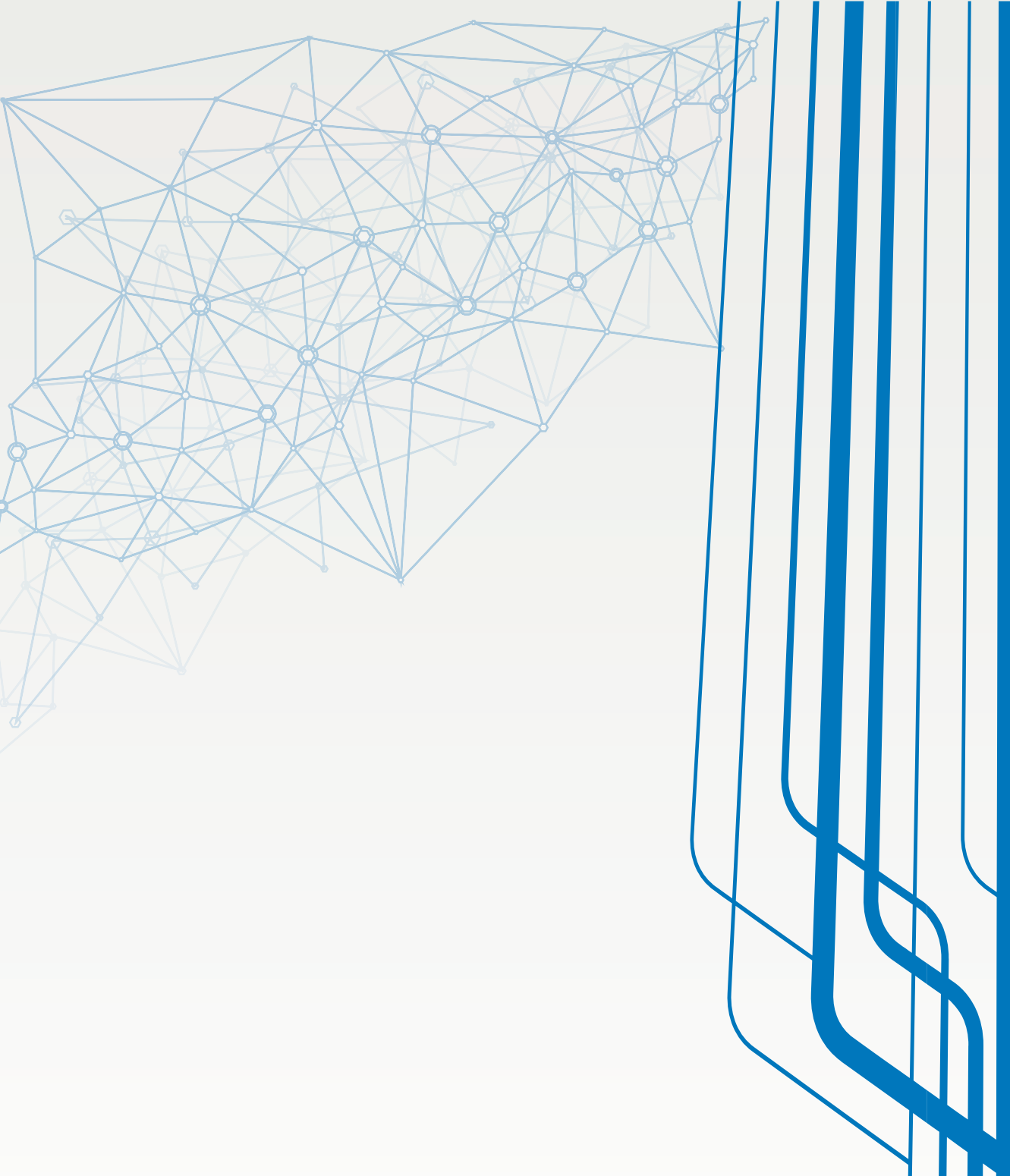
Results show that it is possible to implement security and privacy schemes at IoT devices instead of relying on a single point of failure to enforce those properties for the entire network. It is also important to notice that the presented overhead is regarding an IoT network without any security, which turns security from gateway to cloud irrelevant. To provide means of enforcing user privacy requirements in a cloud environment is paramount for its wide adoption. The Internet of Things amplifies this requirement, since it significantly increases the amount of data sent to the cloud, showing the importance of this work.

Notes









ATMOSPHERE is funded by the European Commission under the Cooperation Programme, Horizon 2020 grant agreement No 777154.

Este projeto é resultante da 4ª Chamada Coordenada BR-UE em Tecnologias da Informação e Comunicação (TIC), anunciada pela Rede Nacional de Ensino e Pesquisa (RNP) e pelo Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), no âmbito do acordo de cooperação Número 51119.