WHITE PAPER AP5
CYBER SECURITY

EXECUTIVE SUMMARY
Cyber security is a main research challenge because 1) citizens' everyday life relies on cyber systems, e.g. energy distribution, and healthcare, and their failures can affect millions of people 2) faults can be caused by people thousands of kilometers far from the affected area 3) one of the main sources of data helping cyber attackers is the people's unaware digital behavior. Thus, a huge effort is needed to prevent, detect and react to cyber-attacks. Moreover, each cyber system often has specific needs, e.g. confidentiality is needed in banking systems, whereas industrial systems put first availability.
The CNR Project Area ***Cyber Security*** fully covers all the above issues by addressing the following topics:

1. ***Cyber-Physical Systems (CPS)*** join security and safety needs, i.e. cyber-attacks may lead to injuries to human beings and loss of lives.
2. **Network security** will investigate several issues as Slow Denial-of-Service (DoS) attacks as well as monitoring TOR (The Onion Router) WEB network for illegal activities.
3. ***Intrusion Detection and Protection*** by means of energy-based security, i.e. the measure of (abnormal) power consumption.
4. ***Privacy*** risk assessment and privacy-by-design methods are needed to guarantee high protection of personal data to enable (big) data analytics.
5. ***Information Sharing and Analytics*** (ISHA): the design of machine learning, artificial intelligence and data analytics techniques able to make sense of large amounts of data.
6. ***Cyber-intelligence on Social Media***: techniques for gathering and analyzing data from Social Media for Intelligence purposes.
7. ***Secure Software Engineering*** assures integrated approaches to face continuous evolution and criticalities rising during all the development cycle of software-intensive systems.
8. ***Access Control and Trust Management*** are among the most important security tools in large distributed systems.
9. ***Cryptography*** is a keyword: reliable, efficient implementations of state-of-the-art algorithms and protocols are required, and must be assessed w.r.t. high-performance code breaking platforms.
10. ***Cloud Security*** concerns the protection of data and resources that are stored and shared on the Cloud, and of the business or research process that are outsourced to the Cloud.
11. ***Cyber insurance*** is a new domain: damages caused by targeted attacks need new mathematical models and regulations w.r.t. accidental events.

The following sections have a sub-section for each of the above topics, in the same order.

1. STATE OF THE ART OF THE RELEVANT SCIENTIFIC AREA

1.1 CPS

W.r.t. the ICT world, Cyber Physical Systems (CPS) have critical security needs due to hw/sw limitations of devices, constraints on power consumption, real-time scheduling and communications, and strongly interconnected security and safety [TII13]. The evolution of CPS towards Industry 4.0 and Factory of the Future exacerbates these needs.

1.2 Network Security

- Slow DoS attacks. Concerning the design of innovative cyber-attacks, relevant contribution to the research world was provided, by introducing the term "slow dos attack", relatively to last generation denial of service attacks, also proving for the first time that such threats can be successfully executed on not performant hosts [IGP16]. Also, innovative protection methodologies, applied to different last generation cyber-threats, were proposed.

- TOR WEB. The exploration and analysis has flourished in the recent past on Web graphs, but not on the TOR network. Little information is available about the topology of the TOR WEB graph and its relation with content semantics [TOR17].

1.3 Intrusion detection and protection

The detection of malware using information hiding, e.g. covert channels, is very hard: the resulting throughput of data for such malicious communication is usually very limited and each covert channel highly depends on how the information is hidden and which hidden data carrier is used. A recent approach relies on high-level indicators to decouple the detection from the underlying technology. In this vein, energy-based security allows to exploit abnormal consumptions to detect information-hiding-capable threats [LC1].

1.4 Privacy

Many practical and impactful services based on big data analytics can be designed in such a way that the quality of results can coexist with high protection of personal data by providing methodologies for privacy risk assessment [PPP+18] and applying a suitable privacy-by-design methodology.

1.5 Information Sharing and Analytics

Information Sharing and Analytics (ISHA) concerns the design of machine learning, artificial intelligence and data analytics techniques able to make sense of large amounts of data. These techniques are particularly effective in identifying system/user anomalies [A+17], and in predicting/preventing security threats and adversarial attacks.

1.6 Social Media

Social media data gathering and analysis for Intelligence purposes, mainly in the fields of hate speech detection, user interaction analysis and face similarity identification; malicious accounts are responsible for manipulating the public opinion [CDP+15] thus requiring techniques for modeling and detecting them by analysing user profiles, posts, and social links.

1.7 Secure Software Engineering

The integration of the proper security management and control during all phases of the development life cycle of software and systems is able to avoid critical security flaws and vulnerabilities. In particular security-by-design is now considered as key solution in different application domain [BCD+17].

1.8 Access Control and Trust Management

Access control and trust management are relevant security mechanisms in large and open distributed systems as internet. This also entails the capability to determine trust levels complex architectures [DMM+18].

1.9 Cryptography

Modern cryptography is a tremendous tool for cybersecurity, and its importance is supposed to increase in the next decades. Three characteristics in the current development and deployment of information services are indeed the *multi-tenancy* of computer environments (as in DaaS, [GS1]), the *multi-authorship* of data (as in blockchain systems), and the *multi-security* for user end points. These three factors are jointly weakening traditional cryptographic approaches, since they break traditional chains of trust, and expose both data and applications to threat conditions which heavily depend on context and can frequently change over time. New cryptographic approaches are required because of emerging cipher-breaking platforms, too. Cost-affordable massive farms and large-scale networks of high performance computing units (eg; GPUs, ASICs) are already used for this purpose [CUB17], and quantum computers are on the way.

1.10 Cloud Security

The ever increasing adoption of the Cloud for executing applications and sharing resources or data, on the one hand gives several performance and cost advantages but, on the other hand, introduces new relevant security issues, being user authentication, even across domains, and authorization for long lasting accesses among the most relevant ones [FGCS16].

1.11 Cyber Insurance

Insurance was only recently applied to the cyber world. The immature cyber insurance market faces a number of unique challenges on the way of its development [AO17].

1.11 Bibliography

[TOR17] Massimo Bernaschi, Alessandro Celestini, Stefano Guarino, and Flavio Lombardi “Exploring and Analyzing the Tor Hidden Services Graph”, ACM Trans. Web 11 (4) pp. 24:1-24:26, 2017

[TII13] M. Cheminod, L. Durante, A. Valenzano, "Review of Security Issues in Industrial Networks," in *IEEE Transactions on Industrial Informatics*, 9(1), pp. 277-293, 2013

[IGP16] E. Cambiaso, G. Papaleo, G. Chiola, M. Aiello “Mobile executions of Slow DoS attacks” *Logic journal of the IGPL*, 24, pp. 1-14, 2016

[LC1] L. Caviglione, M. Gaggero, J.-F. Lalande, W. Mazurczyk, M. Urbanski, “Seeing the Unseen: Revealing Mobile Malware Hidden Communications via Energy Consumption and Artificial Intelligence”, *IEEE Transactions on Information Forensics & Security*, 11(4), pp. 799-810, 2016

[A+17] F. Angiulli, F. Fassetti, G. Manco, L. Palopoli. “Outlying property detection with numerical attributes”, *Data Min. Knowl. Discov.* 31(1), pp. 134-163, 2017

[PPP+18] R. Pellungrini, L. Pappalardo, F. Pratesi, A. Monreale: "A Data Mining Approach to Assess Privacy Risk in Human Mobility Data." *ACM TIST*, 9(3), pp. 1-27, 2018

[CUB17] M. Cianfriglia, S. Guarino, M. Bernaschi, F. Lombardi, “A Novel GPU-Based Implementation of the Cube Attack”. Proc. Applied Cryptography and Network Security (ACNS 2017) ,pp. 184-207, 2017

[CDP+17] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, M. Tesconi, “Social Fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling”, *IEEE Transactions on Dependable and Secure Computing*, 2017

[BDK15] A. Bertolino, S. Daoudagh, D. El Kateb, C. Henard, Y. Le Traon, F. Lonetti, E. Marchetti, T. Mouelhi, M. Papadakis, “Similarity testing for access control”, *Information & Software Technology*, 58, pp. 355-372, 2015

[BCD+17] A. Bertolino A., A. Calabrò, F. Di Giandomenico, G. Lami, F. Lonetti, E. Marchetti, F. Martinelli, Matteucci I.; Mori P., ”A tour of secure software engineering solutions for connected vehicles”, *Software Quality Journal,* pp 1-34, 2017

[DMM+18] Gianluca Dini, Fabio Martinelli, Ilaria Matteucci, Marinella Petrocchi, Andrea Saracino, Daniele Sgandurra: Risk analysis of Android applications: A user-centric solution. Future Generation Comp. Syst. 80: 505-518 (2018)

[FGCS16] E. Carniani, D. D'Arenzo, A. Lazouski, F. Martinelli, P. Mori, “Usage Control on Cloud System”, *Future Generation Computer Systems*, 63(C), pp 37-55, 2016

[AO17] A. Marotta, S. Nanni, A. Orlando, F. Martinelli, A. Yautsiukhin, “Cyber Insurance Survey”, in *Computer Science Review*, 24, pp. 36-61, 2017

[GS1] F. Montecuollo, G. Schmid, R. Tagliaferri, “E2FM: an encrypted and compressed full-text index for collections of genomic sequences”, *Bioinformatics*, 33(18), pp. 2808-2817, 2017

2. CONTRIBUTION TO THE RELEVANT SCIENTIFIC AREA

2.1 CPS

Challenges in securing CPS are dealt with the following research activities 1) description of the CPS by means of a formal model able to cope with all the characteristics of its components, their interactions and the required security policies 2) development of a sw tool able to check whether or not the above model satisfies the security policies 3) study and test how modern communication and service provisioning paradigms (such as SDN and NFV) can improve the security of CPS 4) performance analysis and test of industrial stateful

network firewalls in order to guarantee they can match the high performance required by CPS with real-time constraints, where special purpose communication protocols are used.

2.2 Network Security

2.2.1 Slow DoS attacks

In-depth studies on attack technologies, carried out with the aim of designing innovative defense systems, provided the ability to contribute to relevant research projects. The focus is not only on critical ICT systems in general, but also on specific infrastructures. Also, penetration testing and security tools investigation is accomplished.

2.2.2 TOR WEB

Research activity done in the EU-ISEC IANCIS project has contributed to improve knowledge about TOR Dark Web, providing: a survey over the possible TOR WEB exploration approaches and their limitations; the selection and adoption of a relevant set of metrics for evaluating actual TOR hidden services data; a novel in-depth investigation over the TOR WEB topology; an in-depth analysis of the relationship between the topics found in English TOR pages and the TOR WEB topology.

2.3 Intrusion detection and protection

Challenges in information hiding and network steganography require the investigation of novel detection techniques. Possible new ideas to investigate are: i) statistical / machine-learning techniques able to parse software artifacts (e.g., the distribution of bytecode of executables or the composition or statistics of traffic patterns) to detect the presence of information-hiding-capable threats, and ii) define energy or computational metrics aimed to capture the presence of threats covertly exchange data in emerging scenarios including IoT nodes, high performance computing / parallel systems or network and home appliances. In both cases, scalability issues will be investigated as to allow their implementation to monitor production quality settings or to protect large-scale deployments or wide networks.
Besides, another important challenge deals with energy-aware / green security, which aims at understanding the impact of security on the energy footprint, both for optimization and detection purposes.

2.4 Privacy

Challenges arising from the implications of privacy and data protection issues in ICT world are dealt with the following research activities: 1) designing of data transformations following the privacy-by-design principle for making data private while preserving data quality; 2) designing data mining and big data analytics approaches which by-design guarantee privacy protection; 3) designing privacy-by-design technologies that guarantee *corporate* privacy protection while outsourcing data mining tasks, i.e. the third party cannot infer sensitive information both from data and from extracted knowledge; 4) development of technologies for privacy protection: from data usage control to privacy-aware secure multi-party computation; 5) studying methodologies and tools for software engineering specific for privacy protection; 6) studying methodologies for assessing the privacy risk level by considering different and realistic attack models that are particularly suitable for specific type of data.

2.5 Information Sharing and Analytics

Challenges in ISHA requires the development of techniques for 1) security analytics, based on behavioral profiling, to detect malicious activities and devise models of trust; 2) social sensing for prediction of sensitive information diffusion flows and secure information sharing; 3) attack prevention/response based on machine learning and AI to improve reaction to incidents; 4) privacy-preserving information handling based on theoretically guaranteed models of privacy.

2.6 Social Media

Techniques for gathering relevant amount of data from Social Media; big data analysis for Intelligence purposes in the following fields:

- Malicious accounts detection on Social Media, analysing user behaviour, relations and posted content
- hate speech detection, performing NLP analysis on user posts and comments

- analysis of different types of interactions between users, analysing the evolution of their behaviour considering space and time
- analysing multimedial content, in order to perform face similarity detection on posted images

2.7 Secure Software Engineering

Secure Software Engineering involves different activities in all the phases of development life cycle. They include: 1) secure requirements and quality attributes collection, analysis and design 2) software or system construction, verification, validation and evaluation 3) on-line monitoring, control and assessment of specific security properties and metrics 4) Integration tools and methods for automate security management into all the software process.

2.8 Access Control and Trust Management

Challenges of security mechanisms in large distributed systems promote the research activity in different areas: 1) Verification and Validation (V&V) of access and usage control systems to protect (personal) data and resources against unauthorized, malicious or erroneous usage; 2)Technology transfer of standard V&V approaches in the context of security and privacy; 3) Effective strategies for test case prioritization and selection 4) Test suite effectiveness assessment through mutation analysis;5) Compliance assurance of the rules on international transfers of personal data; 6) Development of model-driven approaches for dynamic access and usage policies specification and evaluation.

2.9 Cryptography

Current main activities in cryptography are as follows. New full-text indices in minute space for performing fast pattern-search queries and evaluation on nucleotide sequence collections in DaaS (Database-as-a-Service) models. The E2FM-index can save about 95% of storage and search for patterns in times of milliseconds. Network architectures and protocols for the management of multi-authorship, distributed databases: functionalities and issues of the promising blockchain technology are under investigation, and a near join work with the DI - University of Milan will explore fairer computing approaches than PoW (Proof-of-Work) and PoS (Proof-of-Stake) systems. On the cipher-breaking side, a smart effective general framework implementing a parallel GPU-based version of the CUBE attack has been designed and implemented. The obtained results allowed to improve the state of the art, managing to attack stream ciphers like Trivium . Other brute force and dictionary-based attacks have been successfully devised and implemented on GPUs, allowing among other things to break the BitLocker encryption system.

2.10 Cloud Security

The research challenges on Cloud security concerns several areas such as: Identity Management and Authentication mechanisms, with particular reference to Federations of Clouds, enhanced mechanisms for Access Control (such as the Usage Control ones) for services, resources and data that shared on the Cloud, virtualization security in multi-tenant environments, data privacy, aspects of compliance, security services (e.g., Policy as a Service), and Cloud governance security. The research activity also takes into account the security issues coming from the integration of Cloud with the Internet of Things and the evolution to Edge computing. The study of the security risks related to the adoption of the Cloud in Public Administration, for example for the management of the citizen's Electronic Health Record is another relevant research topic.

2.11 Cyber Insurance

Cyber security insurance is designed to mitigate losses from a variety of cyber incidents, including data breaches, data theft, business interruption and network damage. Despite a slow start and many problematic issues, the cyber insurance market grows and Insurance companies are increasingly offering such policies, in particular in the USA, but also in Europe. A robust cyber-security insurance market could help to reduce the number of successful cyber-attacks by promoting the adoption of preventive measures in return for more coverage and encouraging the implementation of best practices by basing premiums on an insured level of self-protection.

Considering this topic from a business or an economic perspective too, is strongly required. In light of these considerations, the main contributions in this area concern: studying market solutions for cyber insurance,

improving knowledge on peculiarities of cyber insurance, pricing and risk measures estimation, effect of security interdependence, analysis of cyber insurance as an incentive to invest in security.

3. IMPACT

3.1 CPS

CPS are connected to both the physical world, and ICT systems. Thus, addressing their security needs by a multi-directional approach able cope with all the system components, their relationship and the whole cycle-life, from the design to decommissioning, through all upgrades and updates can improve the way such systems are perceived, also by citizens and users, and contributes to a more secure evolution of these systems, also in an inter-system perspective.

3.2 Network Security

3.2.1 Slow DOS attacks

Concerning cyber-security, the aim is to improve security capabilities for devices, networks and infrastructures to prevent the world from being involved in possible cyber-attacks on individuals or communities.

3.2.2 TOR WEB

A further in-depth analysis of the obtained experimental data helps discovering and describing novel relationships between topology and semantics, and allows a careful reasoning of these results. Indeed, our work shows interesting characteristics of the TOR WEB graph that relate topic semantics and WEB graph topology. Such findings contribute to a better understanding of the TOR usage and of contained information, allowing for more effective Police investigations.

3.3 Intrusion detection and protection

Currently, it is important to investigate new, sophisticated ways of hiding data in communication networks as this allows to detect vulnerabilities in the existing network protocols, which can lead to their malicious exploitation, e.g., by enabling “invisible” confidential data leakage or empower a new-wave of advanced persistent threats. Therefore, this may improve the security of existing protocols and propose some countermeasures to mitigate threats, which can also target transitional scenario such as IPv4/IPv6 ones.

3.4 Privacy

Large datasets recording human activities are key enablers of a new wave of knowledge-based services, as well as of new scientific discoveries. Unfortunately, the use of human data may raise the concern on leakage of personal information. Therefore, organizations need to exploit the advantage analyzing available big data while preventing privacy violations, which may result in negative economic and social impacts. Addressing privacy issues by developing technologies and methodologies for assessing and guaranteeing privacy protection by-design may help in setting the *data free*.

3.5 Information Sharing and Analytics

ISHA for cybersecurity can significantly reduce the risks and effects of attacks, by revolutionizing the way incident response to cyber events is handled. Coupling machine learning and AI techniques with a mathematically guaranteed notion of privacy to augment cybersecurity allows to automatically handle and make sense of complex data flows, detect/prevent attacks, reduce the risks and the effects, and react to breaches while preserving the privacy of data contributors.

3.6 Social Media

Tools have been developed and released to italian LEAs in order to gather data from Social Media and using BigData techniques in order to:

- classify users as malicious/legitimate
- analyze user interactions
- perform face similarity on collected images
- perform hate speech detection on collected messages

- provide tools for performing complex visualizations

3.7 Secure Software Engineering

Secure software engineering is recognized as an effective and efficient mean for: increasing the overall quality level of the developed software or systems; decrease the risk of vulnerabilities and security flaws; drastically reduce the cost and effort for the management and correction security problems; increase the user security perception and confidence in the final products.

3.8 Access Control and Trust Management

Large scale distributed systems and smart environments are requiring complex access control as well as trust management. One the one hand guaranteeing access to resources is one of the paramount aspects of security. In open frameworks, where the environment is not closed, the necessity to monitor the behaviours of entities or of applications is crucial.

3.9 Cryptography

According to the next set of laws by which the EU Commission intend to strengthen and unify data protection for all individuals in the new global digital market, cryptographic algorithms and software will play a major role in the near future. The goals of the action line Cryptography are the design of new cryptographic algorithms and protocols, and the integration of existing state-of-the-art cryptographic tools for the protection of data and processes in emerging digital services and applications. On the cipher-breaking side, obtained results shed a new light on the effectiveness of massively parallel algebraic-based attacks. Further, other successful brute force, dictionary-based attacks we implemented on GPUs can help to show weaknesses in existing, real-word crypto systems.

3.10 Cloud Security

Security is a key aspect of Cloud affecting its adoption by citizens, companies, and also public entities, in particular for what concerns the execution of critical parts of the business/research processes and the storage and sharing of personal, valuable, and critical data. Hence, the design of proper security techniques that enhances Cloud security, for instance by regulating the access and the usage to the resources shared on the Cloud, would have a great impact on the Cloud adoption.

3.11 Cyber Insurance

Cyber insurance by itself provides a unique opportunity to cover risks, as well to contribute to social welfare. Different technological systems impose different challenges on cyber insurance and, at the same time, provide different opportunities. Moreover, a knowledge of cyber insurance is critical because companies need to review cyber security and resilience considering the role of cyber insurance as part of risk management. From insurers perspective, our research can help defining standardized procedures.

4. EMERGING RESEARCH CHALLENGES

Design and analysis methods for CPS security and safety, also matching performance and timing.

Scalable and energy optimized malware detection techniques.

AI and machine learning techniques to analyze large amounts data and preserve users' privacy.

Generation of methods to keep the pace with Social Media evolution for Intelligence purposes.

Health metrics for health ranking of online social ecosystems.

Definition of new approaches and standards for safety and security integration.

Advanced techniques for distributed trust, including Distributed ledgers.

Design and implementation of new cryptographic primitives and protocols, and their security assessment through cipher-breaking methods exploiting new hardware features.

Research to secure the resources and the data shared and processed on the Cloud in the light of the evolution of Cloud to Edge computing.

Methods to overcome interdependent security and information asymmetry issues in Cyber Insurance.

Analysis methods to cope with increasing size of TOR and its evolving graph and content.

5. CONCLUSIONS

The current research activities and proposed approaches are showing good results from the theoretical and practical point of view, with successful publications and collaborations with national and international industrial partners, and public bodies. On the one side this shows that the right way has been taken, but also that more resources and effort are needed to provide adequate solutions in a world where cybersecurity is perceived as a major challenge by both citizens and institutions, and rightly so.

In particular, the use of artificial intelligence and machine learning deserves further investigations: the large and varied amounts of data organization deal with require the design of techniques able to both automatically and precisely detect/prevent anomalous behaviors. In this respect, the AP can play a prominent role and provide both practical and theoretical results.

Research on Social Media Intelligence is producing a strong impact on both scientific research and towards safeguarding from manipulation the information exchanged in our online communities. Emerging challenges need to be faced promptly and strongly, in order to keep pace with the rapidly evolving threats in social media.

Moreover, the different proposed testing solutions adopted in various real word environment have been proven to be effective in detecting security flaws of adopted security systems and improve their overall quality and confidence. Research activity reveals also the necessity of the integration of different knowledge area so to cover the multidisciplinary aspects of large scale and smart environments.

Cryptography is playing a major role because of data and process sharing in modern computing environments. Because of emerging cipher-breaking platforms, it is expected that in the next decade cryptography will be a very active research field.

On the Cyber Insurance side, although it is a desirable option for agents, it has many open issues yet to be resolved by scientists and practitioners.

Novel approaches and treatments are required to ensure the positive effect of cyber insurance on society as well as new standards and practices required for the maturation of the market.

In general, a more and more interconnected world produces and moves huge amounts of data that, on the one side, need to be exchanged, managed, and stored in a secure way, but, on the other hand, these data represent a wider and wider attack surface. For this reason, the future keywords constraints in security will be efficiency, scalability and performance, as, for instance, malware detection techniques require. Missing any of them, will lead to lose the cyber war.

PROJECT AREA 5: CYBER SECURITY

Editorial team and Contact person (CP)	**Institute**	**Email**
AIELLO MAURIZIO	IEIIT	maurizio.aiello@ieiit.cnr.it
CAVIGLIONE LUCA	ISSIA	luca.caviglione@ge.issia.cnr.it
DURANTE LUCA (**CP**)	IEIIT	luca.durante@ieiit.cnr.it
MANCO GIUSEPPE	ICAR	giuseppe.manco@icar.cnr.it
MARCHETTI EDA	ISTI	eda.marchetti@isti.cnr.it
MARTINELLI FABIO (**CP**)	IIT	fabio.martinelli@iit.cnr.it
MONREALE ANNA	ISTI	annamonreale@gmail.com
MORI PAOLO	IIT	paolo.mori@iit.cnr.it
ORLANDO ALBINA	IAC	alba.orlando@cnr.it
SCHMID GIOVANNI	ICAR	giovanni.schmid@icar.cnr.it
TESCONI MAURIZIO	IIT	maurizio.tesconi@iit.cnr.it