Information and Communication Technologies Institute

Carnegie Mellon | PORTUGAL

AN INTERNATIONAL PARTNERSHIP

TERMS OF REFERENCE FOR THE

2009 CALL FOR PROPOSALS

SEPTEMBER 2009

FCT Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

CONTENTS

THEMES FOR THE 2009 CALL FOR PROPOSALS

This document provides relevant information for the Second Call for Proposals by the Fundação para a Ciência e a Tecnologia (FCT) under the Carnegie Mellon | Portugal Program.

Regulations governing access to funding are available at:

http://alfa.fct.mctes.pt/apoios/projectos/regulamento2008

The announcement of the opening of each call is available at:

http://alfa.fct.mctes.pt/apoios/projectos/concursos/cmu/2009

The version of the strategic plan for the Carnegie Mellon | Portugal Program to be considered in the 2009 Projects Call is available at:

http://alfa.fct.mctes.pt/apoios/projectos/concursos/cmu/2009/docs/terms.pdf

In accordance with the strategic plan for the Carnegie Mellon | Portugal Program, which was subject to a broad scientific debate in the research and business communities, research projects should involve scientific institutions in Portugal, companies and research groups at Carnegie Mellon. In this call, priority will be given to the following key focused areas:

- Next Generation Networks for Trusted High-Quality Services
- Software Engineering for Large-Scale Dependable Systems
- Cyber-Physical Systems for Ambient Intelligence

Projects are expected to meet the objectives of the aforementioned strategic plan, to be inter-disciplinary in nature and to combine fundamental research with the development of new technologies. High-quality projects in other areas pertaining information and communication technologies may also be funded.

Duration of activity. A typical proposal should cover research efforts for a period of three years. Funding for years 2 and 3 will be conditional on the progress achieved by the team as determined upon review by the Board of the Carnegie Mellon | Portugal program.

Funding. The number of funded projects will be limited. The total funding for the beneficiaries from research units and universities in Portugal is estimated at a total of 2.4 Million Euros. The participating

researchers from Carnegie Mellon University will be funded by the Information and Communication Technologies Institute at Carnegie Mellon University. The involved companies will not be funded by FCT.

Eligible applicants. Each proposal must be submitted by a consortium of at least two research groups associated with different Universities, Associate Laboratories, Research Centers, State Laboratories or any other public or private research institutions in Portugal, at least one company and at least one research team from Carnegie Mellon University in a thematic area relevant to the call. It is expected that students enrolled in the dual degree MS and PhD programs of the Carnegie Mellon |Portugal program, as well as students of the first cycle under the Bologna declaration, be integrated in the research activities carried out at the research units and the involved companies. Beneficiaries must have a recognized curriculum vitae in the areas relevant to the execution of the project. Priority will be given to proposals that include the effective participation of industrial partners, materialized in the form of human resources at companies and the commitment of financial or material resources.

Each researcher with a PhD can only be involved in at most two different proposals.

Where to apply. Applications must be submitted online through the FCT website https://concursos.fct.mctes.pt/projectos following the Announcement of the Call for Proposals.

Deadline for receipt of full applications. The call is open from September 7, 2009 until November 13, 2009 at 17:00 Lisbon time.

Requirements. Full applications must fulfill the following requirements:

- Be written in English
- Include an overview of proposed research emphasizing how the proposal meets the objectives of the Carnegie Mellon | Portugal Program. Emphasize also the nature and expected outcomes of the collaborative work. State clearly the aims of the team as a whole and indicate the specific role played by each member in

achieving the goals of the team. Explain the advantages of conducting the proposed studies as a collaborative effort. (do not exceed three pages, Arial regular font, 11 point)

- Include a detailed outline of the proposed research, including preliminary results (do not exceed seven pages; Arial regular font, 11 point). The detailed presentation must delineate the specific contributions of each member of the team.
- Include the CV and publication list of the principal applicant and each co-applicant (for each applicant, maximum of 2 pages, Arial regular font, 11 point, including a list of up to 10 of the most relevant papers)

Review of applications. Applications will be reviewed by international panels of independent experts, as organized by FCT. Review panels will be responsible for evaluating the merit of each proposal.

The selection for funding is based on the following criteria:

1. Relevance of the project towards obtaining comparative advantages for Portugal, in accordance with the objectives stated in the strategic plan for the Carnegie Mellon | Portugal program and the terms of reference defined for each thematic area;
2. Scientific merit, relevance, originality, inter-disciplinarity and expected results of the proposed research;
3. Scientific merit, scientific productivity and collaboration potential of the research groups and their qualifications to carry out the project;
4. Ability to integrate young PhDs, PhD students and MS students (involved or to be involved in the Carnegie Mellon | Portugal program), as well as undergraduate students in the work to be carried out at the research units and the involved companies;
5. Degree of involvement of companies in the proposed research and transfer of the research outcomes to the industry;

6. Rationale for the proposed research and budget.

Notification. Successful applicants will be notified by e-mail after February 1, 2009.

Start of activity. Funded applications will commence after March 1, 2010.

Progress Report. A yearly report will be delivered, for interim review by the Board of Directors. The Board has the right to recommend that further funding be discontinued or extended for one more year.

Final Report. A final report will be delivered, for review by the Board of Directors.

KEY STRATEGIC AREAS

NEXT GENERATION NETWORKS FOR TRUSTED HIGH-QUALITY SERVICES

As many of the world's communication and computation infrastructures expand, connect and evolve into one pervasive mesh of heterogeneous devices, we witness large investments to bring fiber directly to most homes and the fast convergence of large-scale data processing facilities, ultra-portable computers, intelligent vehicles, safety and health-care systems, energy-aware technologies and peer-to-peer architectures. Network operators, equipment manufacturers and government regulators are now confronted with ever changing system architectures, traffic demands, and costumer expectations with respect to quality of experience, brought now to new levels for example by high-definition television over IP. Not satisfied by simply acquiring information from centralized servers, many of today's users want to upload their own content and share it with the world. Likewise, businesses are becoming increasingly global with teams collaborating across continents and time zones, requiring massive data transfers and high-speed connections for multimedia communication. The high-performance networks of the future must provide their customers with unhindered access to the world wide web of people and things, which must be available at all places and in all moments, automatically and without delays, irrespective of the underlying optical, cable or wireless infrastructure, yet automatically adapting to and taking optimal advantage of the available features of the access device.

MAIN TECHNICAL CHALLENGES:

- **Design and Integration of Future Internet Technologies:** Seamless communication and service continuity over the emergent global network requires innovative technologies, which can be integrated, optimized and operated with ease over multiple domains, from the physical communications channel up to the services and applications as seen by the user. This includes an all fiber infrastructure capable of carrying massive traffic across the network to individual homes and businesses, as well as an array of wireless gateways that enable mobile services and reliable connectivity in dense and volatile environments. The finite amount of wireless spectrum asks for dynamic resource allocation based on new hardware and software for flexible radio. It is important to investigate how novel lower layer techniques ranging from cognitive radio to multiple-antenna systems and cooperative communications (e.g. relaying, beam forming, and network coding) can be leveraged by the communications protocols at higher layers of the system architecture to provide the necessary throughput and robustness guarantees. Mesh networks, home networks and vehicular networks will complement the cellular infrastructure in ways that are yet to be explored, with users roaming from one network to another seamlessly or even tuning simultaneously to multiple access points for higher rates and efficiency. Decisions on the best radio access technology selection may be based on context awareness and localization. The selfishness of users and devices that compete for common network resources can be taken into account based on game-theoretic models and the search for equilibrium. Increasingly the network will interact more and more with the physical world through wireless enabled objects, also known as the Internet of Things, which are likely to change traffic patterns and alter services and applications thus justifying the development of new architectures and protocols.

- **Security, Privacy and Trust:** The growth of the ICT sector is strongly dependent on the level of confidence with which the average costumer is willing to carry out electronic transactions over the network. It is not sufficient for the infrastructure, protocols and services to be secure, their level of security must be obvious to anyone wishing to use them. This includes the correct usage of enabling technologies for security like

cryptography, protection against spoofing attacks on wireless and cable links, active measures against identity theft by means of phishing, virtual machines to isolate malware, trust primitives for software updates, traitor tracing for peer-to-peer content, intrusion detection, and other defense mechanisms capable of adapting to the constant surge of new attacks and vulnerabilities. Electronic IDs can be included in the effort towards implementing reliable identity management. Security should not be an add-on to existing communication protocols and services, but rather be a primary concern in the design of new information and communication technologies.

- **Applied Machine Learning for Traffic analysis and Efficient Network Management:** Peer-to-peer communication and social networking are only two examples of recent developments that are fundamentally changing how information and data packets flow within a large-scale communications network. By collecting and compressing massive quantities of data on network traffic and employing tailored machine learning and data mining techniques, it should be possible to characterize, at least partially, the resulting patterns of network usage. A natural step could be to investigate how root-cause analysis can predict demands, anticipate problems, reduce congestion and combat failures. The key is to understand how the decisions made by a myriad of devices based on local rules and partial information impacts the emergent behavior of large-scale networks as a whole.

- **Network Assessment and Evaluation:** How customers value the services provided by high-performance networks is ultimately determined by the levels of satisfaction they reach during their interaction with the various systems. Standard metrics such as packet loss rates or transmission delay are clearly insufficient to capture the end user’s satisfaction, which is closely tied to human perception and our ability to process intelligible speech or tolerate certain classes of visual artifacts. Consequently there is need for research on testing methodologies and measurable criteria by which networks can be assessed and compared. More adequate metrics should be incorporated into the optimization and planning that governs the interaction between different operators and user terminals. The objective must be to maximize network efficiency whilst minimizing the effort of each network entity and guaranteeing a prescribed level of quality of experience. Multimedia services

should be context-aware and adapt to the conditions of the network and the terminal of the user. Network virtualization, with large pools of networked physical machines that can host tens of thousands of virtual machines and services, emerges as a possible tool to provide differentiated services requirement specific methods for management and assessment.

- **Regulating, Charging and Billing:** With the paradigm shifting from connection oriented traffic to distributed systems with constant roaming among multiple service providers, it is not yet clear how governments should regulate the telecommunications sector and how businesses can charge for their services in a fair and effective manner without incurring excessive management overhead. Economic trends are pointing towards a fragmented market, in which communication services are often offered by autonomous service providers who are independent of the customer management, access network provision and core network management entities. Load balancing techniques must also be revisited to account for new modes of communication in which users are as much service providers as service recipients. Ultimately, we must seek ways for the network infrastructure and its protocol stack to provide the means to track individual users without compromising their privacy, measure their contributions and their benefits, and produce a billing scheme that ensures the economic viability of successful services.

SOFTWARE ENGINEERING FOR LARGE-SCALE DEPENDABLE SYSTEMS

Software is a key enabler of economic development as it provides the means to store, process and exchange information as varied as catalogs of products, items in a warehouse, positioning of vehicles, market transactions, and energy measurements. To meet the challenges of globalized software development, where thousands of components must be combined to provide the customer with an adequate application or information system, program modules must be engineered to operate in a seamless and reliable way with each other, adapting to unpredictable scenarios, recovering from unexpected breakdowns and providing simple means to maintain or extend their features and configurations. To satisfy these requirements, software cannot be developed in an ad-hoc fashion. Thus, better software engineering methodologies and tools offering agility and fast delivery emerge as an indispensable tool for developing the information systems of the future.

MAIN TECHNICAL CHALLENGES:

- **Dependable Software Systems:** The development of industrial strength software requires new ways to certify the integrity, the correctness and the inter-operability of different modules and systems, even when such subsystems must operate across large-scale networks, and subject to continuous updates and extensions. This objective can be attained by means of formal methods from computational logic and type theory, which can be used to develop new high-level design and implementation idioms that are intrinsically reliable and extensible, as well as verification tools. Some intervention by human operators is likely to be necessary even in the presence of self-healing, self-adaptive software. Therefore, research should be devoted to identifying critical scenarios in which human oversight and actions are indispensable, leading to novel methodologies for dealing with highly complex software systems.

- **Collaborative Workflow Software:** Many software packages from business applications to multi-player games are now being developed by large teams that are spread out in various parts of the globe. This is possible because workflow software and online collaboration tools allow for a project to be divided into multiple chunks that are sent out to different developers, whose work is then assembled to produce the final product. New infrastructures for distributed software development will provide the fertile ground over which global innovation networks are likely to grow.

- **Software Modules for Global Supply Chains:** Inventory tracking, automatic transactions and consistent distributed databases are some of the key ingredients of modern supply chain management. Software systems for multinational enterprises must enable reliable real-time monitoring of goods and processes, thus ensuring that vital data is freshly available to all key partners and decision makers at all stages of the global supply chain. Research in this direction must rely on the real-world data collected by internationally operating companies. Challenges also include dealing with shared data-sets where parties with conflicting interests recognize the benefits of collaboration, but impose strong bounds on the admissible information flows.

- **Multi-core and Parallel Systems:** The multi-core processor is today a standard building block. It is commonly used in the edge (servers, proxies) and in the core (routers switches using network processors) and it is an appropriate computational substrate for telecommunication workloads because this class of workloads tends to exhibit large amounts of parallelism. Sharing of the computational capacity of multi-cores contributes heavily to the end-to-end delay that users experience. A comprehensive theory with algorithms for sharing resources and for proving upper bounds on the delay experienced by an individual program does exist, however this theory is arguably not yet well developed for multi-cores. Moreover, it does not take into account the effects of contention for internal buses and switched interconnection networks inside a multi-core. It is now clear that future massive parallel processing will bring concurrency to the center of software engineering concerns. Harvesting the benefits of parallelization will require new programming abstractions and tools for expressing parallel programs at a high level, verifying correctness properties of concurrent code, and efficiently executing code on modern multi-core platforms.

CYBER-PHYSICAL SYSTEMS FOR AMBIENT INTELLIGENCE

The convergence of computation, communication, sensing and control capabilities in relatively small devices at very low cost has signaled the advent of so called cyber physical systems, composed by large numbers of independent nodes that take measurements from physical processes and interact with the surrounding environment. These massively distributed systems can have many different applications ranging from the safety monitoring of roads, bridges, buildings, water distribution systems and public transportation systems to the provision of remote healthcare services and support to first responder in emergency scenarios. To face the challenges posed by numerous application scenarios, sensors and actuators must coordinate their efforts to achieve a global behavior that mirrors ambient intelligence, i.e. the ability of cyber-physical systems to use the collected data in a distributed fashion to solve problems in a collaborative way.

MAIN TECHNICAL CHALLENGES

- **Robust Sensor-Actuator Nodes:** The ability of a cyber-physical system to meet the challenge of a particular

application clearly depends on the characteristics of individual nodes, which must be small enough to be embedded in the environment, robust enough to survive adverse conditions, tamper-proof to resist malicious attacks, flexible enough to adapt to new conditions or applications, and complex enough to be able to sense, compute and communicate over a typical noisy channel. In particular, new sensing and radio technologies are necessary, which can operate for example inside large buildings or in underground structures.

- **Distributed Sensing and Actuation:** To enable applications to run on cyber-physical systems, signal processing, communication and control algorithms must be re-thought to account for the strong restrictions in terms of power and computational capabilities. The information picked up by heterogeneous sensors must be secured and can be combined with other information about the physical world that is generated by users and made available through the web (e.g. pictures posts with GPS information). Node mobility, for example through network robot systems, changes the picture once again, forcing cooperative control to work even in the presence of highly volatile channels. Design issues include cooperative localization, navigation, environment perception, map building, task allocation, and task execution.

- **Networked Infrastructures:** As sensor nodes are embedded seamlessly in buildings, water pipes, electrical grids and even inside textile fabric, these resulting cyber-physical systems will be able to collect very large amounts of data, compute the necessary inferences, and make decisions that can increase the life-cycle of these infrastructures, identify possible risks and prevent disruption of service or attacks on their integrity.

- **Middleware for Large-Scale Deployments:** The design of massively distributed applications poses great challenges to the programmer and to developers of programming languages. A top-down design would allow for the system designer to specify the functionalities of the system as a whole, which could then be translated semi-automatically to code that runs on each individual node. These micro-programs and their updates and configurations then have to be spread over the network, which from an operational point of view also requires dedicated middleware and adequate interfaces for the system administrator. Embedded nodes are becoming more and more

powerful, but complex to specify and develop. In one hand, new multiprocessor and multi-core technologies are being used for small devices, which allows building more powerful applications, but that imposes new requirements in terms of resource management (CPU, memory, power, etc). On the other hand, applications become more complex to specify, imposing the need for advanced, but more and more reliable, operating systems and programming languages.

- **Scalable Query Processing:** The large number of sensing devices produces an enormous amount of sensor readings and clearly this generates an enormous amount of traffic. Fortunately, application designers are not interested in obtaining every sensor reading but designers are rather interested in answering high-level queries such as “What is the maximum temperature in this area?” or “Is there a person in this room?”. It is therefore typically not necessary to communicate all sensor readings since it is possible to compute answers to such queries inside the network. Novel approaches have been recently proposed that excel one of the main features of CPS: tight coupling of computing and communications with the physical environment and dynamics.