EXPERIMENTAL COGNITIVE DISTRIBUTED ENGINE

ECODE

Scope

The goal of the ECODE experimental project is to develop, validate, and implement a cognitive routing system, which uses machine learning techniques, that can meet the future Internet challenges, in terms of manageability, security, availability, and accountability. This project combines networking with machine learning expertise, in order to address these challenges in the most efficient way.

During the building phase, the cognitive routing system is both designed and prototyped. Then, in the second phase, the project considers three sets of use cases to evaluate the benefits of machine learning techniques. The project will develop, implement, and experiment new techniques to improve:

- "Manageability and Security" of the Internet: efficiently monitor the path performance by combining passive and active measurements, and cooperatively detect traffic anomalies (leading to performance or QoS decrease) so as to detect intrusions and attacks;
- "Availability and Accountability" of Internet paths: efficient path ranking based on QoS and availability metrics, efficient path re-routing to other links, in cases of failure, and traffic flows correlation by routers in order to diagnose and predict their deviation over time (with respect to profile-based resource allocation), and adapt these profiles so as to maintain an acceptable resource usage;
- "Scalability and Quality" of the Internet routing system: by detecting events that are detrimental to the routing system dynamics and to efficiently react to such events.

For this purpose, the project will investigate novel semi-supervised, on-line, and distributed machine learning techniques kernel of the cognitive routing system. The experimentation and the validation of the techniques developed in the project will be carried out on physical (iLAB) and virtual (e.g., OneLab) experimental facilities.

Technical and innovation approach

The driving idea of the ECODE project is to augment the existing Internet system and network lower-level data

Type of project
Small/medium-scale focused research project (STREP)

Contract number
FP7 – 223936

Project coordinator
Dimitri Papadimitriou

Contact person
Dimitri Papadimitriou
Alcatel-Lucent Bell
Copernicuslaan 50
B-2018 Antwerpen
Belgium

Tel: +32-3-2408491
Fax: +32-3-2404886
dimitri.papadimitriou@alcatel-lucent.be

Project website
www.ecode-project.eu

Community contribution to the project
3.070.000 Euro

Project start date
1 September 2008

Duration
36 months

collection and decision making, with a cognitive component. This component enables the Internet infrastructure to learn about its own behaviour and environment over time in order to better analyse problems, to tune its operation, and to enforce its decisions on manageability, security, availability, and accountability, so as a result to better satisfy end-users. Thus, the introduction of a cognitive component provides the lower level mechanisms and means by which the Internet would resolve any new challenges that result from its evolution: a growing end-user basis with higher heterogeneity in their needs, and a wider utilization for which the Internet was not initially designed.

The ECODE project focuses on five main challenges, linked to the evolution of the Internet. By combining the fields of networking research field with the machine-based research, the ECODE project proposes to redesign the control capabilities of the IP/networking layer, in an unprecedented way. Indeed, no satisfactory solution is currently available to address these challenges altogether. More precisely, two classes of challenges are addressed by this project:

- Operational challenges: ***manageability/diagnosability***, and ***availability,*** caused by a performance drop, due to an increasingly complex and growing Internet infrastructure, for which existing solutions are no longer adequate. In particular, availability problems result from the decreasing routing system quality (in particular, its stability, its robustness, and its convergence properties);
- New challenges: ***security*** is only supported weakly by the current Internet infrastructure. On the other hand, ***accountability*** (part of the initial Internet design objectives) has never been really met by the current Internet infrastructure, e.g., traceback and other congestion control techniques are not widely deployed to identify misbehaving users and traffic sources. As the Internet grows, the ***routing system scalability*** progressively results into major cost concerns for both vendors and Internet Service Providers (ISP).

Our methodology relies on cross-fertilization between the networking and machine-based domains to form a cognitive routing system answering the operational and new Internet challenges. Indeed, they are similar in nature to the conditions traditionally encountered in classical machine learning problems:

- *Nature*: the events cannot be well characterized even when examples of such an event are available (inherent complexity in precisely characterizing an event);
- *Relationship*: the correlations and trends between events are hidden within large amounts of data that are associated to these events;
- *Environment*: the conditions are changing over time (this is particularly the case for the routing environment but also variability of user demands, expectations and behaviours);
- *Quantity*: the amount of available data is too large for handling by human intervention;
- *Evolutive*: new events are constantly detected/discovered.

From this analogy, the main concept of this project is to extend existing IP networking equipment, with a distributed cognitive engine (as shown in Fig.1) based on semi-supervised, on-line, and distributed machine learning techniques. The cognitive engine derives a number of observations from the data collected from the routing and forwarding engine and other cognitive engines. From these observations, the cognitive engine learns rules that result in local decisions (at the forwarding and routing level) and distributes them to other cognitive engines. The distribution of the processing as well as the learned rules may depend on the peering relationship between cognitive routers. In particular, the boundary defined by autonomous systems may be a limiting factor to the distribution of such information.

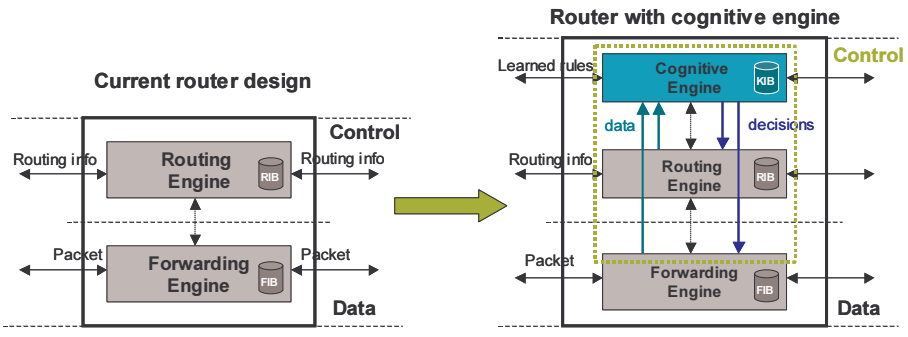

Fig.1: Cognitive engine and Cognitive router

The Future Internet Research and Experimentation (FIRE) initiative launched by the European Commission has put forward a method to experimentally investigate Future

Internet technologies. The FIRE initiative acknowledges three main characteristics of experimentally driven long-term research: openness of the approach, multi-disciplinarity, and research by experimentation and practice. In this context, the ECODE project aims at the creation, the development and the experimentation of a distributed cognitive engine (based on semi-supervised, on-line, and distributed learning techniques) that coexists and sits next to the existing routing engine of current IP networking equipment. This project experiments the capability of this cognitive component to solve the operational and new Internet challenges. Like the application of graph theory in the networking domain has allowed for the emergence of the well-known routing discipline, the introduction of a distributed cognitive component cooperating with the routing and forwarding engines allows them to evolve beyond their current capabilities and give rise to a new networking paradigm for the Internet.

Target users and benefits

By introducing a new architectural component, the ECODE project is expected to 1) improve and extend the Internet functionality[1] by providing for the adequate solutions to the existing and foreseeable upcoming Internet challenges, 2) limit the cost of the Internet infrastructure growth, and 3) reduce the operational cost and complexity of the Internet (compared to the approach that would consist in continuously patching existing routing equipment). This project also targets the durability of the Internet (and the valid design principles of its current architecture) by removing the complexity, from existing components, but adding functionality.

The results of this experimental project will increase Europe's strength as a pioneer in the multidisciplinary research field of networking and machine- learning research domains along with the standardization actions at the Internet Research Task Force (IRTF) and Internet Engineering Task Force (IETF) including:

- The open communication protocols between routers' cognitive engines.
- The interfaces between the cognitive engine and the existing components of IP routers;
- The cognitive routing system functionality, its architecture, and its components.

The ECODE project results could be exploited in two different ways:

- Novel IP routing equipment, which is able to sustain the Internet evolution, growth, demands and loads. The project should result in a "revisited loosely coupled control infrastructure" of the Internet able to cope with its long-term evolution and growth. Indeed, the Internet infrastructure would be easier to manage and able to operate under increasingly challenging conditions, and would offer a higher level of availability;
- Application of advances in machine learning: semi-supervised learning is still a partially open issue. Here, the full prediction problem would be based on partially labelled training data only. Another outcome of this project would be new on-line learning algorithms, as the necessarily limited amount of past data to be stored in each cognitive engine would stress the on-line nature of the learning. Finally, whereas some distributed learning algorithms exist, the possibility of distributing the estimation procedure of such models would be of high value.

Project partners

Organisation name	Country
Alcatel-Lucent Bell	BE
Universite Catholique de Louvain	BE
Universite de Liege	BE
Interdisciplinair Instituut voor Breedbandtechnologie	BE
Institut National de Recherche en Informatique et Automatique	FR
Lancaster University	UK
Centre National de la Recherche Scientifique	FR

[1] in terms of functionality (manageability, security, availability, and accountability) and performance (in terms of scalability)