International Journal of Engineering Education Vol. 32, No. 4, pp. 1566–1578, 2016
Printed in Great Britain
0949-149X/91 $3.00+0.00

Enhancing Engineering Education through Link Prediction in Social Networks*

ÁNGEL FERREIRA-SANTIAGO and CORNELIO YÁÑEZ-MÁRQUEZ
Instituto Politécnico Nacional—Centro de Investigación en Computación IPN-CIC, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal, CIC, Mexico City 07738, Mexico. E-mail: cyanez@cic.ipn.mx

ITZAMÁ LÓPEZ-YÁÑEZ, OSCAR CAMACHO-NIETO and MARIO ALDAPE-PÉREZ
Instituto Politécnico Nacional—Centro de Innovación y Desarrollo Tecnológico en Cómputo IPN-CIDETEC, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal, CIDETEC, Mexico City 07738, Mexico. E-mail: ilopezy@ipn.mx, oscarc@cic.ipn.mx, maldape@gmail.com

AMADEO-JOSÉ ARGÜELLES-CRUZ
Instituto Politécnico Nacional—Centro de Investigación en Computación IPN-CIC, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal, CIC, Mexico City 07738, Mexico. E-mail: jamadeo@cic.ipn.mx

In recent years the world has been a witness to a brutal onslaught of emergent technologies. As such it is not surprising that social networking has permeated through practically every human activity with amazing speed. Educational systems have not lagged behind; and not only is that true, but it is also evident that social networks have ostensibly penetrated in engineering education. This relevant and irrefutable fact has generated the necessity of posing systematic research activities on a worldwide scale that are related to this topic. In this context the topic covered in this paper is Social Network Research related to Engineering Education. Specifically, our research concerns the way or ways in which link prediction in social networks is able to improve teaching-learning processes in engineering education. A suitably natural environment for the application of such research is that of scholarly publications related to computer science, particularly networks. The results of our research are promising; they facilitate valuable information regarding the tendencies of the actors immersed in engineering education, be they students or faculty members.

Keywords: engineering education; social networks; link prediction; computer science publications.

1. Introduction

Technology attracts an enormous amount of interest. In recent years the world has been a witness to a massive surge of emergent technologies. A significant amount of material has been written about the ways technology is currently changing our lives, about the possibilities of its development in the future and of the hopes that it will revolutionize the way people learn and exchange information. The advent of connected computers means that the road is paved for the emergence of enhanced and engaged learning activities because of the ability to not only bridge the physical gap between people scattered across enormous distances, but to facilitate autodidacticism; that is, an instructor is not always needed for learning in this system. This means that a student can now enrich his or her learning experience by exchanging ideas and content across multiple areas of knowledge via distributed environments [1].

Another purpose of networking technology is to construct a shared learning space, both on a local and global scale. The needs of students are now greater: not every learner comes from the same sociocultural background. Not every learner possesses the same skill set, nor do they have identical prior knowledge of the subjects. Learning can be enhanced when these factors are taken into account, and as such teachers should keep this in mind when designing the learning environment. The hope is that emerging technologies boost this learning process in ways never seen before. However, the mere existence of a technology by itself cannot guarantee more success in learning unless it is utilized properly; as such, the way technology is employed in the learning process is the key factor that impacts online teaching and learning [2]. The way to produce an effective learning tool is to study the human brain and how cognitive processes work, in order to produce quality educational software which truly enhances learning [3].

Technology applied to learning has permeated the teaching-learning landscape so much that specialists now fiercely argue about its role: some praise emerging technologies as a triumph of modernity which helps society, while others dismiss technology as a misused and potentially abusable innovation [4]. Setting aside this debate, and considering the enormous influence of emergent technologies applied to educational environments, the fact that social networking has permeated through practi-

* Accepted 15 February 2016.

cally every human activity—and done so with amazing speed—is not surprising. In fact, educational systems not only have not lagged behind in this regard, but it is also evident that social networks have ostensibly penetrated in the way education is engineered.

A community seen as a social network consists of members reciprocally interacting with each other. These interactions produce cohesion between members while being inaccessible to outsiders. These networks, being based on computers, do not limit membership to people within the same location neither require everyone to be connected at the same time. Members in a network can also provide other resources to each other: information, feedback, advice, job opportunities and news, among others. Thus, educational networks can prove to be a friendlier environment of communication between peers [5].

Social networks can be modelled as nodes, consisting of people, linked by edges as connections between those individuals; in other words, a graph. The influence a person has on a social network is determined by his connectivity: the number of people to whom he or she is linked and the number of paths they form throughout the network. Assessing the impact of an individual within the social network has several applications: for example, data mining can extract patterns to determine where cliques are forming, measuring reputations and monitoring social happenings. However, this observation can be performed at various levels, since people demonstrate a tendency to interact more closely and with higher frequency with people with whom they share interests or opinions [6].

The author of a recent work, *Mining the Social Web* [7] affirms that social networking is indeed changing the landscape of our lives, both online and offline, and that they induce various effects in us, enabling both constructive and destructive behavior in its users. The Internet is trending toward ubiquity more and more every day, and social networking is just one step in this process. Data mining, analysis and visualization techniques are among the tools to answer questions involving certain actions between actors in a social network; for example, the composition of a person social circle, the way people communicate with their peers and the degree of reciprocity between friends, among others.

The authors of the present paper consider that searching for answers to the questions regarding usage habits in social networking like those mentioned above is important. We are convinced that there exists a need for researching the role of social networking in education on a worldwide scale. Specifically, our research concerns the way or ways in which link prediction in social networks is able to improve the teaching-learning process in education engineering. We assert that a suitably natural environment for the application of such research is that of scholarly publications in the field of computer science.

The rest of the paper is organized as follows: some basic concepts about social networks as well as link mining and link prediction are described in section 2, while the context and discussion which make up the main contribution of this paper are presented in section 3. Section 4 is dedicated to conclusions and future work, while bibliographical references and the author biographies are included at the end.

2. Presentation

This section is dedicated to describing the social networks as well as basic concepts about link mining and link prediction, focusing on the main goal of the present paper: link mining in social networks. At the end of the section, some recent research works related to these topics are presented.

2.1 Social networks

Members of a social network interact with each other; by establishing relationships inside educational environments, these networks may become a much more effective medium of communication among peers. Emerging social networks bear social capital as inherent value, manifested in the common interests that help foster near instant contact between social network members with affinity [5].

Graphs are an adequate tool to model social networks; nodes represent people and edges are connections between individuals. Personal computers and mobile devices such as smartphones and RFID or Bluetooth tags allow connection to social networks which offer interaction and communication. Social networks can be characterized by the diameter, distribution and degree of its nodes or by the mean connectivity of the nodes; leading to the identification of various roles in the nodes. An important task in social networks is finding key actors in function of their connectivity, the number of contacts they have and the paths that traverse each node. The influence a person has on a social network is determined by his connectivity: the number of people to whom he or she is linked and the number of paths they form throughout the network. Assessing the impact of an individual within the social network has several applications: for example, data mining can extract patterns to determine where cliques are forming, measuring reputations and monitoring social happenings [6].

As social networks develop, thanks in great part to the advancement of world wide web-related technologies, groups of researchers worldwide have entered this field, taking measurements and collecting data. These efforts are geared to applying the scientific method to the study and understanding of the relationships between individuals immersed in social networks.

Formally, a social network is defined as a finite set of social actors (nodes). These nodes, which are no more than the network's members, are connected by one or more kinds of relationships (edges), whose characterization is of great help to researchers. Concepts exist which substantiate the measurements and characterizations of networks, such as ties, density, centrality, cliques and other relevant features [8].

Links—In a network's representation as a graph, links (also known as edges or ties) are connections between two or more of the network's nodes. Links can be directed or undirected. Directed links can also be unidirectional or reciprocal, and they can be modeled as binary relationships or can be weighted so that some ties are stronger and others are weaker. A clear example of a directed link in the educational context is when one of the members (a student) sends his work to another member (the teacher). An example of a reciprocal directed link is that of two students sharing ideas for solving a problem related to a course that both are taking. An example of weighted links would be the frequency with which two specific students exchange ideas; in some cases, this weight would be almost zero, when those two seldom interact.

Density—This concept describes the general linkage level in a social network's graph representation. In an informal way we can say density measures how far is a network's graph from being complete; in quantitative terms, density is the cardinality of the link set divided by the cardinality of the set of vertices of a complete graph with the same number of nodes.

Path—Nodes or actors of a social network can be directly connected by edges or ties, or indirectly by a sequence of links. A trail is a sequence of lines in a graph, and a path is a trail in which each point and each line are different.

Length—The number of lines that make up a path is defined as its length.

Distance—The distance between two nodes is the length of the shortest path which connects both nodes.

Centrality—The most active nodes or actors are detected by the measure of the concept called centrality; meaning the nodes which relate to other members with the most intensity and reach. There are three important measurements regarding the concept of centrality: (a) Degree centrality: means popularity or intense activity; it is the sum of all the nodes connected to a specific node; (b) Between-ness centrality: the capacity of a given node of connecting pairs of other nodes; that is, in a way it is the measure of the node's potential to function as a controller of the resource flow in the network. This measure indicates how powerful is that node in the network as a whole; and (c) Closeness centrality: measures efficiency and independence of a specific node. If path lengths from a certain node to others are lesser that two, it is said that that node can reach any other node by itself and does not depend on other nodes to reach any desired node.

Clique—Given a graph which represents a social network, a clique is a subgraph in which any of its nodes is directly connected to any other node in the subgraph.

In the proposal presented in this work, these concepts are used while emphasizing the concept of link within the structure of social networks in the environment of engineering education.

2.2 Link mining and link prediction

Links in a social network possess a conceptual wealth such that they exhibit relevant properties for researchers; among these properties we can find the category, importance or rank of the nodes and of the links between them when they are considered as objects of study. In some cases it is not possible to observe all of the links; hence, it can be interesting to predict the existence and weight of links between specific nodes; in other cases of social networks where links evolve over time, it could be of interest to predict whether a link will materialize in the future [7, 9].

According to [9], the expression "Link Mining" has been coined to represent all data mining techniques that explicitly consider nodes and links in a social network as objects of study. Typically there are eight tasks in link mining that focus on objects, links and graphs:

Link-based object ranking—Based on exploiting the link structure in a graph representing a social network with the goal of assigning priorities to the nodes in the graph. The objective of this task is to arrange the nodes in order of importance according to the measure of their degree centrality.

Link-based object classification—This link mining task consists of labeling nodes by categories, according to the structure and characteristics of its links. In contrast to traditional pattern classification, this task faces the challenge of the labels being possibly correlated, which adds increased difficulty when designing classification algorithms.

Object clustering (group detection)—Clustering of nodes has as its purpose detecting groups in the social network that share common properties. The challenge faced with this task is to make contributions in research related to the processes of knowledge discovery.

Object identification (entity resolution)—The purpose of this task is to determine which references point to the same node inside a social network. Entity resolution has been considered as a pairwise resolution problem, where each pair of references is resolved independently. In link mining it is desired to use links to improve entity resolution, and for that reason one has to take into consideration, along with the attributes of the references to be resolved, the rest of the references with links pointing to their nodes.

Link prediction—The task of link prediction is of central importance in the present paper; it consists of predicting the existence of a link between two nodes based on the attributes of the nodes and of other links in a social network. Typically, the researcher observes some links and based on the gathered information predicts links that have not been yet observed; however, the most common way to describe the task of link prediction from the temporal viewpoint is as follows: the state of a set of links is known at the time t and the state of that link set at the time t+1 is predicted.

Clique discovery—Given a social network, by tending to the task of clique discovery one attempts to discover cliques of interest with common properties, or that conform to a certain specified pattern.

Graph classification—Differing from the link-based object classification task, which consists in labeling the nodes by category, the task of graph classification is a supervised learning problem in which the goal is to categorize a whole graph as an instance of a concept, be it positive or negative.

Generative models for graphs—This task attempts to discover general principles that dictate the different kinds of networks. Recently, general patterns that lead into generative models for graphs have been observed when studying, among others, the structural properties of biological and social networks.

Link prediction stands out among the eight main tasks of link mining due to its reach and relevance in social networking for an educational environment. The authors of the present work have posed the question of what is the way in which we can improve engineering education using link prediction in social networks, which have demonstrated the ability to be effective academic helpers in the teaching-learning processes in engineering.

Link prediction is evidently a particular case of link mining. The purpose of this work is to establish the importance of link prediction in engineering education. With a social network within the context of an engineering education environment, the task of link prediction can be described as a binary classification problem; that is: let n_i and n_j be two different nodes of a social network, which can be potentially linked; the problem to solve is to predict if the link l_{ij} exists or in other words, whether it is equal to zero or one.

Several approaches exist to take on link prediction; among these, we can mention measurements of graph proximity, regression models, probabilistic and statistical models, random Markov fields and relational representations.

2.3 Link prediction in social networks: notation and basic concepts

The content of the present section is strongly based on references [10–12]. The task of link prediction is quite relevant in the field of social network analysis, although it has also relevant applications to other domains such as information retrieval, bioinformatics, and e-commerce. By applying such a link prediction algorithm, several tasks are enabled, for instance: identifying spurious interactions, extracting missing information, or evaluating network evolving mechanisms. For example, in social networks currently non-existent but very likely connections may be discovered, and thus recommended as promising friendships, which is a feature that can be helpful for users in finding new friends, improving the user experience, and even strengthening their loyalty to the social network. Another example is the application of similar techniques to survey the way in which a particular social network (or portion of it) grows and evolves in time.

In the context of social networks, links between nodes usually represent connections or friendships between users. Thus, the manner in which these links evolve is commonly driven by mutual interests between users, which are intrinsic to the group to which they belong. However, social networks are inherently very fluid and dynamic, with new links and nodes added, as well as removed over time. Understanding the dynamics underlying the evolution of social networks (or even portions of them) is a complex problem, due to the large number of variable parameters that must be considered. In this regard, a sensible simplification is to study the connection between two particular nodes. Thus, some questions of interest that may arise are: What factors influence the dynamics of such links? In particular, how is the link between two nodes affected by other nodes? In general, how does the link pattern evolve over time?

Given a social network, the convention is to denote the sets of nodes and links by N and L, respectively, while $x \in N$, $y \in N$, and $z \in N$ denote specific nodes at time t. Also, $\Gamma(x)$ represents the set of neighboring nodes of $x \in N$. Based on this notation, the task of link prediction may be defined as follows. Let a social network be represented by a graph $G(N, L)$, where N denotes the set of nodes and L the set of links. Thus, a link between nodes $x \in N$ and $y \in N$ at some time t is denoted as $l(x, y)$. If T represents a time after t (i.e. t<T), then the clique of G containing the set on links existant between times t and T is represented by $G[t, T]$. Now, let $[t_1, T_1]$ be a training time interval and $[t_2, T_2]$ be a testing time interval, where $T_1 < t_2$. Then, the task of link prediction consists of finding the set of links not present in the clique of $G[t_1, T_1]$, but which will appear in the clique of $G[t_2, T_2]$.

Using this notation, it is possible to state the task of link prediction as a supervised pattern classification problem; for this, the feature space is considered to be the set of all pairs of nodes (x, y) in graph $G(N, L)$. The training set is formed by all pairs or nodes (x_1, y_1) in the training interval $[t_1, T_1]$, which are labeled as follows: the pairs of nodes in $[t_1, T_1]$ such that $l(x_1, y_1) \in L$ are given the label 1, while the pairs of nodes in $[t_1, T_1]$ such that $l(x_1, y_1) \notin L$ are labeled with 0. Now, for the supervised classification phase, let (x_2, y_2) be a pair of nodes in the testing time interval $[t_2, T_2]$; then, the task of link prediction consists of predicting the label of $l(x_2, y_2)$.

Current scientific literature includes several approaches to link prediction, as formulated previously. For instance, there are probabilistic algorithms, such as relational models or those based on the Bayesian decision theory. On the other hand, some methods use a kernel matrix and involve maximum margin classifiers, while there is another group based on maximum likelihood. Also, some algorithms are based on graph evolution models while others make use of linear algebraic formulations.

Among this wide range of algorithms, a group based on similarity score calculation between two nodes with the goal of later applying a supervised learning algorithm stands out. In this context, and with the goal of generating the patterns features, there exists a set of effective methods for measuring proximity, taken from ideas in the fields of graph theory, computer science and the social sciences, which will be outlined below. First we describe the indices based on node neighborhoods:

Common neighbors—This index reports the number of neighbors that any two nodes, $x \in N$, $y \in N$, share. It is the cardinality of the intersection between the sets containing the neighbors of both nodes:

$$|\Gamma(x) \cap \Gamma(y)|$$

Jaccard coefficient—This index normalizes the common neighbors index, and is expressed as follows:

$$\frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|}$$

Salton index—This index, also called Cosine Similarity Index, is defined as follows:

$$\frac{|\Gamma(x) \cap \Gamma(y)|}{\sqrt{|\Gamma(x)| \cdot |\Gamma(y)|}}$$

Leicht-Holme-Newman index—This index assigns higher similarities to node pairs with several common neighbors, comparing this number not with the maximum possible number of neighbors, but with the expected quantity of neighbors. This index is defined as:

$$\frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x)| \cdot |\Gamma(y)|}$$

Sørensen index—This index is defined as follows:

$$\frac{2|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x)| + |\Gamma(y)|}$$

Hub promoter index— Used for calculating topological features and is defined as:

$$\frac{|\Gamma(x) \cap \Gamma(y)|}{min\{|\Gamma(x)|, |\Gamma(y)|\}}$$

Hub depressed index— It is a variant of the Hub Promoter Index and is defined as follows:

$$\frac{|\Gamma(x) \cap \Gamma(y)|}{max\{|\Gamma(x)|, |\Gamma(y)|\}}$$

Resource allocation index— This index is used for measuring the amount of resources that a node receives from another in a social network and is defined as:

$$\sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{|\Gamma(z)|}$$

Adamic-Adar index—This index assigns higher weights to the least-connected neighbors; it has been used for measuring similitude between two web pages. It is defined as:

$$\sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{log|\Gamma(z)|}$$

Preferential attachment index— The underlying premise of this index's definition is in the evolution of a social network: the probability that a new link has x as a final node is proportional to $|\Gamma(x)|$; under this scheme, the probability that a new link connects any two nodes $x \in N$, $y \in N$, is proportional to:

$$|\Gamma(x) \cap \Gamma(y)|$$

There exist as well methods for measuring similarity score based on paths.

Katz index— This index takes into account the sum of all paths that exist between two nodes $x \in l$, N, and is defined as follows:

$$\sum_{t=1}^{\infty} \beta^t . |\text{paths}_{(x,y)}^{\langle t \rangle}|$$

where β is a positive parameter, t represents a path length, and $|\text{paths}_{(x,y)}^{\langle t \rangle}|$ is the set of all paths of length t that join the nodes x and y.

Hitting time—This measure indicates the expected value of the number of steps a random walk would take to traverse the path from x to y. That is, in a social network, the hitting time $H_{x,y}$ is the expected amount of steps in a path starting at node x which iteratively moves to some neighbor of the last visited node, which is chosen uniformly at random from the set $\Gamma(x)$ until it arrives to the node y.

SimRank—It is a recursive process in which a parameter called *decay factor*, defined as $\gamma \in [0, 1]$, intervenes. For two nodes x, y the SimRank indicator is defined so that $SimRank(x, x) = 1$ and $SimRank(x, y)$ is:

$$\gamma \cdot \frac{\sum_{a \in \Gamma(x)} \sum_{b \in \Gamma(y)} SimRank(a, b)}{|\Gamma(x)| \cdot |\Gamma(y)|}$$

The indices described above are not the only ones in the literature, but they are indeed the most important and enjoy the highest usage in scientific works related to link prediction.

2.4 Link prediction in social networks: recent relevant work

In this section, recent works related to link prediction on social networks are discussed. In [13], Murata and Moriyasu use structural properties of online social networks in order to do link prediction in Question-Answering Bulletin Boards, which can support a recommender system for potential answerers. On the other hand, Liu and Lü tackle link prediction with methods based on local random walks: in [14] they explain two particularly difficult aspects of link prediction on social networks, which are the sparsity and huge size of the targeted networks.

The roles that both weak ties and node centrality play in link prediction on social networks are two topics which have been studied by multiple authors. In [15] Lü and Zhou estimate the likelihood of the existence of links in weighted networks by means of local similarity indices, such as the common neighbors, Adamic-Adar index, the resource allocation index, and their weighted versions. These authors conclude that the resource allocation index performs best, once they applied their methods for information retrieval. Also, Liu, Hu, Haddadi, and Tian [16] base their proposal on node centrality and weak ties, in order to achieve hidden link prediction on five real-world networks; for this, each common neighbor plays a different role for the node connection likelihood depending on their centralities.

The authors of [17] argue that the straightforward, standard method for performance evaluation of missing links prediction may lead to terrible bias, proposing instead that missing links are more likely to be links connecting low-degree nodes. Thus, they analyze how to uncover missing links with low-degree nodes, namely links in the probe set are of lower degree products than a random sampling. These authors test their proposal on four different real networks, further introducing a parameter-independent index which considerably improves the prediction accuracy.

The topic of social network topology information for link prediction is broached in [18] by Dong, Li, Yin, and Rui. These authors claim that the existing link prediction algorithms do not apply the network topology information sufficiently, and thus propose two improved link prediction algorithms: one is based on the global information network, while the other is based on multiple attributes of nodes. For verification, the proposal was tested on a DBLP data set

The paradigm of unsupervised learning is also present in social networks link prediction. In [19] Feng, Zhao, and Xu use clustering to induce the relationship between the social network structure and the performance of link prediction methods. These authors found, by experimenting ob both synthetic and real-world networks, that as the size of the clusters grows, the prediction methods precision could be remarkably improved. On the other hand, the prediction performance is poor for sparse and weakly clustered networks.

The authors of [20] consider a cluster in graphs to be a group of vertices which is densely connected within itself, while being sparsely connected to other groups of vertices. Also, they claim that such clustering results contain the essential information for link prediction, and that common neighbors to some vertices will play different roles depending on whether they belong to the same cluster or not. The authors propose a link prediction method based on clustering and global information, which they test on data sets taken from synthetic and real-world social networks.

Recently, an evolutionary approach to link prediction on dynamic social networks was introduced in [21]. There, the authors provide a method to predict future links by applying the Covariance Matrix Adaptation Evolution Strategy to optimize the weights of a linear combination of node similarity and neighborhood indices. Then, they applied successfully this method to a large dynamic social network of more than a million nodes: Twitter reciprocal reply networks.

The reference [22] is a very interesting work, in which an ordered weighted averaging (OWA) operator based link prediction ensemble for social networks is presented. The authors begin by assuming that local information-based algorithms have high variabilities, thus proposing a stable link predictor with low variance. This is a link prediction ensemble algorithm based on OWA operators, for social networks; this novel method assigns aggregation weights for nine local information-based link prediction algorithms with three different OWA operators. The advantage of doing so is that, when this technique is applied to benchmark social network datasets, the experimental results are more stable.

3. Context and discussion

In the Alpha-Beta Group (ABG) we are committed to instruction and cutting-edge scientific research. Along more than a decade of existence, the members of the group have experienced increasing activity concerning an important common goal: Enhancing Engineering Education (EEE) at an undergraduate as well as postgraduate level; that is, without neglecting the cornerstone of the activities of the ABG, which is precisely high level scientific research in current topics.

With the purpose of contextualizing our incipient incursion into social networks, it is necessary to mention advances in the storage and retrieval of concept lattices [23], where the models that the ABG has cultivated from its beginnings, Associative Models (AM), are successfully applied. These models along with Neural Networks and their applications have been present in the scientific activities of the ABG, and in several occasions our research activities are inspired by the accomplishments obtained by other research groups in topics as relevant as Hopfield neural networks (which can also operate as associative memories), to cite just one example [24].

Sustainable development is also present in the research activities of the ABG. The study of atmospheric pollutants has been conducted from the viewpoint of time series; that is, pollution concentration measurements (especially in Mexico City) are expressed as data points in a time series and original models, stemming from AM and designed at the ABG, are applied. One notable example of this is the Gamma classifier, which has become one of the flagship successes of the ABG [25].

In 2012 we strongly linked this topic with EEE, with the publication of an article where a novel solution to the secure exchange of environmental education data was proposed [26]; this original solution was heavily inspired by two sources: on one hand, the seminal ideas included in the doctoral dissertation of one of the authors, who is a member of the ABG; and on the other hand, the concepts conveyed by Lytras in a document that has already become a classic in the field of sustainable development research [27]: an editorial pertaining to a special edition about Information Systems Research for a Sustainable Knowledge Society, published in 2010.

Topics with social impact, especially those related to health, are not foreign to the interests of the ABG; on the contrary, for the past three years activity in these topics has been intense. In [28] an associative memory approach to medical decision support systems is proposed; in that very same year (2012) we related this topic in a close way with EEE by publishing a new tool for engineering education: hepatitis diagnosis using Associative Memories [29]. Recently, in 2014, we published a scientific work which includes original models created at the ABG, with a look to applying intelligent algorithms in the field of health: the one-hot vector hybrid associative classifier for medical data classification [30].

In the field of computer science an area of great relevance is Software Engineering, especially the topics related to prediction of the behavior of software development teams, regarding the effort invested in project development or prospective development duration. The ABG has successfully applied its foremost development, the Gamma classifier, to development effort prediction of software projects [31]; currently we are in the process of diving into large scale projects via identification and quantitative analysis of project success factors, inspired by [32].

Tasks related to the algorithms that underlie the workings of wireless networks also motivate study in the ABG. In 2013 we published two works regarding this topic. In one of them we applied the theory of binary decision diagrams (BDD), which served as basis for the design of an algorithm for the minimum spanning tree in wireless ad hoc network routing [33]; additionally, contributions were made to the state of the art in fast route convergence in dynamic power controlled routing for wireless ad-hoc networks [34]

During the present year, 2014, the participation of the ABG in several state of the art topics has been abundant, as well as in topics of EEE, which is one of the group's conceptual cornerstones. We have continued working on the creation, generation, design, development, implementation, and application of AM to areas of great interest and relevance for contemporary Mexican society. For instance, in [35] we published a novel associative model for time series data mining, where the data are intimately related to naturally fractured oil wells prospective in the Mexican state of Tabasco.

If scientific contributions to the state of the art are relevant for the ABG, the topics related to EEE are also one of its priorities, as evidenced by the publications on this hot topic presented during 2014. In [36] we tackle a topic related to emerging technologies. There, we parted from some of our pedagogical experiences in order to highlight the impact that emerging computational tools have on EEE, especially on the field of computer science learning. In this paper we describe, among other aspects, how day to day mentors, teachers, researchers, tutors, scientists, and pedagogues witness the ubiquity of these tools. Also, the manners in which these emerging computational tools have revolutionized the scope of educative and research processes are emphasized. The conclusions of this paper point that classrooms, research and learning labs, field works, and interdisciplinary sessions with collaborative study evidence the breadth and depth of novel didactic resources which those in charge of educating new generations can take advantage of, both for engineering education in general, and computer science learning in particular.

Several authors have been influential in an impactful way on the interest that the ABG shows in EEE topics; with the goal of exemplifying the kind of concepts and ideas that serve as support in our research, we will say that the influence of [36] in our works lies on the vision that the author shows related to framing technology enhanced learning environments, and the notions and perspectives that are present in his paper.

Aside from this, the singular framework for knowledge management in higher education using social networking, presented in [38], has been very useful. It is precisely this work, in concordance with the concepts expressed in [39], which incentivated the members of the ABG to delve into a new research area: social networks, specifically the topic of link prediction.

The interesting concepts and ideas included in [39] which awoke interest among the members of the ABG, are frame in a special edition whose topic is information and communication technologies for human capital development; in this context, the authors affirm that in the last decade, emerging technologies have evolved and as a result formulated a new environment for using human behavior as a basis for knowledge and learning intensive settings. They stat that they have found six technologies with a marked influence in the development of contemporary scientific research; among these six technologies, the one related to social and wireless networks plays a highly important role. Amidst the topics related to this technology we can mention: the formulation of social networks, link prediction in social and complex networks, the rewarding motivational scheme for active participation, the micro-content dimension and the enabling technologies. The strongest motivation stems from the colophon of this work, where the authors state: "As a result of the special issue we would like to highlight some of the most interesting research areas on the forthcoming years on this topics inviting researchers from all around the world to contribute to the scientific debate...", and as the first item of research they nominate the topic of "Knowledge Acquisition in complex structured knowledge, learning and social networks".

The effervescence caused among the members of the ABG by this sort of motivation had the role of conceptual platform in the process that allowed our incipient incursion into social networks. Then, a stage of intensive preparation in the topics described in subsection 2.1, with relation to the representation of a social network as a graph and its associated concepts: links, density, path, length, distance, centrality and clique.

Immersed in this scenario, the decision was taken to work in link mining, and there is where first ideas were set into motion with the goal of identifying the topic that garnered the most interest; among the ones presented in subsection 2.2: link-based object ranking, link-based object classification, object clustering, object identification, link prediction, clique discovery, graph classification, or generative models for graphs. In the end the decision was made to dive into the task of link prediction in virtue of its reach as well as its relevance in social networking for an educational environment.

Social networks throughout their history have demonstrated their worth as helpful aides in academic tasks, such as the teaching-learning process in engineering. Therefore, the question this work aims to answer is: can link prediction in social networks better engineering education and, if so, how?

From this point forward, and arduous process of wide-ranging and deep documental investigation was initiated, among prestigious journals worldwide on the state of the art of link prediction in social networks. For this purpose several technical aspects were taken into account, such as the indices described in subsection 2.3: common neighbors, Jaccard coefficient, Salton, Leicht-Holme-Newman, Sørensen, hub promoter and depressed, resource allocation, Adamic-Adar, preferential attachment, Katz, hitting time and SimRank.

Along with technical aspects, pertinent questions were posed regarding the type of data and the appropriate social networks for doing a study on link prediction. In particular, aspects such as the following were discussed: What are the important characteristics of activity in social networks that allow them to have an active role in EEE? What kind of specific tasks in link prediction could be useful for EEE's purposes? What topics are of interest for the ABG as well as for EEE processes? What kind of data are available in the repositories at hand for researchers worldwide? Are there available data sets for the ABG's research related to link prediction for EEE?

The findings of this arduous research were extremely relevant. The authors of [10] use five coauthorship networks G, obtained from author lists of articles from two sources: arXiv (www.arxiv.org) and CiteSeer (www.citeseer.com), with the goal of performing link prediction; in this context, the task of link prediction consists of predicting which authors will be coauthors in the future. This indicates the importance of social networks that join, by means of coauthorships (links), authors of scientific papers (nodes). The role that this kind of social networks can play in EEE is evident.

In [13] data from the Chiebukuro Japanese Yahoo! Answers (http://chiebukuro.yahoo.co.jp) are used. Here, communications on QABB connect the users, and the social network emerges from the overall connections among users. If it is indeed the case that changes in a social network can be predicted, the usefulness is evident, for communication among users can be encouraged. As an example of an application, link prediction on QABB can result in recommending questions to people who might be able to answer.

There exists an important repository for social networking called "Pajek datasets" (http://vlado.fmf.uni-lj.si/). From there the authors of [15] selected two datasets: USAir, the US air transportation network, consisting of flights (links) connecting two airports (nodes); and CGScience, consisting of a co-authorship network for computational geometry publications, in which a link is a co-authorship of a paper or book joining two researchers, who act as nodes. In addition, in this paper C. elegans is used, the neural network of the nematode worm C. elegans. Here, a link joins two neurons, or nodes, if a synapse or gap junction connects them [40].

The topic of scientific papers coauthor social networks appears again in [18], where the authors use experimental data sets taken from the Digital Bibliography and Library Project (DBLP, http://dblp.uni-trier.de/) to test the performance of the link prediction algorithm.

The work presented in [20] introduces something new: the authors use five synthetic social networks of their creation along with clustering models. Also, five typical real-world social networks data sets taken from different domains were used to validate their method: a) Karate is a social network of interactions between members of a karate club; b) Jazz is a network of jazz bands in which a link between two bands is established if they had at least one musician in common; c) PG is a well-connected electrical power grid of western US, where nodes denote generators, transformers and substations and links denote the transmission lines between them; d) PB is a social network of US political blogs; and e) from the Pajek data sets repository, the authors chose one of the social networks used in [15], which is the US air transportation network, where a link is a flight between tqo airports (the nodes), and is known as USAir.

The work in [21] presents the application of link prediction on Twitter reciprocal reply networks (RRNs). Here, the authors examine the evolution of such networks constructed a time scale of weeks, where nodes represent users and links represent evidence of reciprocated replies during the time period of analysis.

The Pajek datasets repository is quite popular; the authors of [22] picked up four social networks from this repository to test their OWA operator for the task of link prediction: World Soccer Data Paris 1998-WSDP98, Food Webs-ChesLower, and Graph Drawing Contests Data-C96 and B97.

The use of online social networks is reported in [41], whose authors collected the graph data from the Facebook and Hi5 web sites at two different times, using a webcrawler. Moreover, for comparison purposes they used the Epinions data set, which is a who-trusts-whom social network that consist of positive and negative edges (here a positive edge implies trust whereas a negative edge implies distrust).

In the recent publication [42], the authors combine three topics of interest for the ABG: link, prediction, the medical research domain, as well as a recurring topic in the present work which has gained the attention of the ABG, which is co-authorship social networks. The authors tested their proposals on co-authorship networks in the research field of coronary artery disease, taking the data from the Web of Science (WoS) site, consisting of rich information for publications, including authors, publications, titles, references and so on.

Despite the broad diversity of social networks discussed in the various papers considered in our study, there is one topic which exhibits a preponderant recurrence: scientific co-authorship social networks, undoubtedly a very useful field for EEE processes. Inside this topic, the range of social networks from which the corresponding data sets of authors data (nodes) and co-authorship relationships (links) includes such web sites as arXiv, CiteSeer, CGScience, DBLP, and the Web of Science.

However, during the extensive documentary review done by the members of the ABG, a good portion of papers include the study of link prediction on a scientific co-authorship social network known as NetScience. This network is introduced in [44], and is also included in the Pajek datasets repository. A sample of publications consulted in the current paper which make use of the Sciencenet social network for the task of link prediction includes: [11, 14, 16, 17, 19, 43].

The scientific co-authorship social network NetScience is a community from which arises a network of coauthorships between scientists working and publishing in the area of theoretical and experimental networking. The network has a total of 1589 scientists in it, from a broad variety of fields. It is also noteworthy that all of the highlighted authors are group leaders or senior researchers of groups working in this area.

It is the opinion of the current paper authors that applying the task of link prediction to NetScience is undoubtedly of great import to the actions related to EEE become effective. In this sense and given the relevance that high level scientific publications have acquired in relation to engineering education, the ABG is convinced that applying link prediction to the NetScience social network is a clearly effective path to improving the teaching-learning processes related to the educative contents of engineering areas, particularly those of computer and networking sciences (both theory and practice).

As a closing remark for this study in which we have highlighted the improvements which can be made in engineering education through link prediction of scientific co-authorship social networks such as NetScience, we will illustrate some results obtained with several different methods.

According to [11], among the experiment of link prediction on social networks the K-fold cross validation technique has become practically a de facto standard. In this validation method, the set of observed links is randomly partitioned into K subsets of equal size. Then, each time one subset is elected as test set, the rest K-1 subsets make up the training set. The cross-validation process is then repeated K times, with each of the K subsets used exactly once as the test set. As a results, all links are used for both training and validation, and each link is used for prediction exactly once. Notice that a large value for K leads to smaller statistical bias, but requires mode computational resources. Experimental evidence suggests that 10-fold cross validation is a very good trade-off between cost and performance. Also, there is an extreme case where the number of folds (K) is equal to the number of nodes: the leave-one-out method. The latter is usually suggested for small sized data sets.

Among the standard metrics frequently used to quantify the accuracy of prediction algorithms, there is one of particular importance: the area under the receiver operating characteristic curve (AUC), which evaluates the algorithm's performance according to the whole list of links to predict. Provided the rank of all non-observed links is known, the AUC value can be interpreted as the probability that a randomly chosen missing link is given a higher score than a randomly chosen nonexistent link. If all the scores are generated from an independent and identical distribution, the AUC value should be about 0.5. Therefore, the degree to which the value exceeds 0.5 indicates how much better the algorithm performs than pure chance.

In [11], the authors experimented with the task of link prediction on the NetScience social network. They tested, among others, the common neighbors, resource allocation, and SimRank indices, as well as their proposed method based on Random Walk. When taking the mean on 1000 implementations, the best AUC value is that of their method: 0.993, closely followed by SimRank with 0.992. The results presented in [14] are similar.

The work proposed in [16] is based on node centrality and weak ties, and has as goal to achieve hidden link prediction on five real-world networks, including NetScience. For this, each common neighbor plays a different role for the node connection likelihood depending on their centralities. When using degree centrality in the common neighbors index (DC-CN), the authors obtained their best AUC value: 0.938. Even though this value is inferior to those achieved in [11], this result is still very good, since this method outputs link prediction with a high degree of accuracy, according to AUC.

The authors of [17] analyze how to uncover missing links with low-degree nodes, namely links in the probe set are of lower degree products than a random sampling. Experimental analysis on ten local similarity indices, obtaining each value by averaging over 100 independently implementations on NetScience reveals a surprising result: that the Leicht-Holme-Newman index performs the best with 0.992. The 10 indices used are: common neighbors, Jaccard coefficient, Salton, Leicht-

Holme-Newman, Sørensen, hub promoter and depressed, resource allocation, Adamic-Adar, and preferential attachment.

In [19] Feng, Zhao, and Xu use clustering to induce the relationship between the social network structure and the performance of link prediction methods. These authors found, when applying clustering to a data set taken from NetScience, the resource allocation index gives the best results.

The scientific co-authorship social network NetScience is also used for link prediction in [43]. There, the authors examine the link prediction task from the novel perspective of network OWS. They show the existing dependency between the manner in which how easily two nodes interact with each other or influence each other, and both their positions in the network as well as the nature of the OW that mediates interactions between them. Additionally, they measure the performance if different heuristics on link prediction over the NetScience social network, showing that the heuristic based on a random walk-typo process outperforms the popular Adamic-Adar method.

Arbitrated and indexed journal publications have been becoming increasingly relevant to scientific communities around the world, particularly publications included in the JCR index. Mexico is not an exception to this behavior, situation that is evidenced by the fact that both Conacyt and the IPN (institution which hosts the ABG) require their researchers to publish in such journals. This is where such social networks as NetScience gain importance, in the measure in which researchers are capable of doing link prediction tasks, in order to appreciably improve the teaching-learning processes in engineering.

4. Conclusions and future work

In this paper, several research works related to link mining and link prediction in social networks have been discussed. In particular, the potentially positive impact of social networks link prediction on improving engineering education has been explored and argued. As a result, the benefits of solving the task of predicting future links in a scientific co-authorship social network (such as NetScience) has been identified and proposed as a relevant path of improvement to enhance teaching-learning processes in engineering education.

Derived from the present work, the members of the Alpha-Beta Group intend to improve and extend currently existing methods for link mining, incorporating emerging methods and concepts such as the use of metaheuristics to fine tune the parameters of associative models. In this regard, data sets taken from social networks such as NetScience or from repositories like the Pajek data set repository become the starting points for our future work in these research lines.

Acknowledgements—The authors would like to thank the Instituto Politécnico Nacional (Secretaría Académica, COFAA, SIP – grant 20144197, CIC, CIDETEC), the CONACyT, and SNI for their economical support to develop this work.

References

1. M. D. Lytras and P. O. de Pablos, Software Technologies in Knowledge Society, *Journal of Universal Computer Science*, **17**(9), 2011, pp. 1219–1221.
2. L. Zhuhadar, R. Yang and M. D. Lytras, The impact of Social Multimedia Systems on cyberlearners, *Computers in Human Behavior*, **29**(2), 2013, pp. 378–385.
3. M. Vargas-Vera and M. D. Lytras, Exploiting semantic web and ontologies for personalised learning services: towards semantic web-enabled learning portals for real learning experiences, International *Journal of Knowledge and Learning*, **4**(1), 2008, pp. 1–17.
4. C. Yáñez-Márquez, M. Aldape-Pérez, I. López-Yáñez and O. Camacho-Nieto, Emerging Computational Tools: Impact on Engineering Education and Computer Science Learning, *International Journal of Engineering Education*, **30**(3), 2014, pp. 533–542.
5. N. Koudas, G. Vossen and M. D. Lytras, Editorial: revisiting the (machine) semantic web: the missing layers for the human semantic web, *IEEE Transactions on Knowledge and Data Engineering*, **19**(2), 2007, pp. 145–148.
6. M. D. Lytras and P. O. de Pablos, *Social Web evolution: integrating semantic applications and Web 2.0 technologies*, IGI Global Information Science Reference, New York, 2009, p. xvii
7. M. A. Russell, *Mining the Social Web*, O'Reilly Media, Inc., Sebastopol, CA, 2011, pp. 313–319.
8. P. O. de Pablos and M. D. Lytras, Editorial: Semantic Web and Social Networks for E-learning, *Journal of Web Engineering*, **12**(5), 2013, pp. 361–362.
9. L. Getoor and C.P. Diehl, Link mining: a survey, *ACM SIGKDD Explorations Newsletter*, **7**(2), 2005, pp. 3-12.
10. D. Liben-Nowell and J. Kleinberg, The Link Prediction Problem for Social Networks, *Journal of the American Society for Information Science and Technology*, **58**(7), 2007, pp. 1019–1031.
11. L. Lü and T. Zhou, Link prediction in complex networks: A survey, *Physica A: Statistical Mechanics and its Applications*, **390**(6), 2011, pp. 1150–1170.
12. M. Al Hasan and M. J. Zaki, A survey of link prediction in social networks, in CC Aggarwal (ed.), *Social network data analytics*, Springer US, 2011, pp. 243–275.
13. T. Murata and S. Moriyasu, Link prediction based on structural properties of online social networks, *New Generation Computing*, **26**(3), 2008, pp. 245–257.
14. W. Liu and L. Lü, Link prediction based on local random walk, *Europhysics Letters*, **89**(5), 2010, p. 58007.
15. L. Lü and T. Zhou, Link prediction in weighted networks: The role of weak ties, *Europhysics Letters*, **89**(1), 2010, p. 18001.
16. H. Liu, Z. Hu, H. Haddadi and H. L. Tian, Hidden link prediction based on node centrality and weak ties, *Europhysics Letters*, **101**(1), 2013, p. 18004.
17. Z. Yu-Xiao, L. Lü, Q.-M. Zhang and T. Zhou, Uncovering missing links with cold ends, *Physica A: Statistical Mechanics and its Applications*, **391**(22), 2012, pp. 5769–5778.
18. L. Dong, Y. Li, H. Yin, H. Le and M. Rui, The Algorithm of Link Prediction on Social Network, *Mathematical Problems in Engineering*, **2013**, Article ID 125123, 2013.
19. X. Feng, J.C. Zhao and K. Xu, Link prediction in complex networks: a clustering perspective, *The European Physical Journal B-Condensed Matter and Complex Systems*, **85**(1), 2012, pp. 1–9.

20. F. Li, J. He, G. Huang, Y. Zhang and Y. Shi, A Clustering-based Link Prediction Method in Social Networks, *Procedia Computer Science*, **29**, 2014, pp. 432–442.
21. C. A. Bliss, M. R. Frank, C. M. Danforth and P. S. Dodds, An evolutionary algorithm approach to link prediction in dynamic social networks, *Journal of Computational Science*, **5**, 2014, pp. 750–764.
22. Y.-l. He, J. N. K. Liu, Y.-x. Hu and X.-z. Wang, OWA operator based link prediction ensemble for social network, *Expert Systems with Applications*, **42**, 2015, pp. 21–50.
23. M. E. Acevedo, C. Yáñez-Márquez, and M. A. Acevedo, Associative Models for Storing and Retrieving Concept Lattices, *EURASIP Mathematical Problems in Engineering*, **2010**, 2010, Article ID 356029.
24. R. Sammouda, N. Adgaba, A. Touir and A. Al-Ghamdi, Agriculture satellite image segmentation using a modified artificial Hopfield neural network, *Computers in Human Behavior*, **30**, 2014, p. 436–441.
25. I. López-Yáñez, A. J. Argüelles-Cruz, O. Camacho-Nieto and C. Yáñez-Márquez, Pollutants time series prediction using the Gamma classifier, *International Journal of Computational Intelligence Systems*, **4**(4), 2011, pp. 680–711.
26. R. Sepúlveda-Lima, C. Yáñez-Márquez, I. López-Yáñez and O. Camacho-Nieto, A Novel Solution to the Secure Exchange of Environmental Engineering Education Data, *International Journal of Engineering Education*, **28**(6), 2012, pp. 1380–1387.
27. M. D. Lytras, From the Special Issue Editor: Information Systems Research for a Sustainable Knowledge Society, *Information Systems Management*, **27**(3), 2010, pp. 196–197.
28. M. Aldape-Pérez, C. Yáñez-Márquez, O. Camacho-Nieto and A.J. Argüelles-Cruz, An associative memory approach to medical decision support systems, *Computer Methods and Programs in Biomedicine*, **103**(3), 2012, pp. 287–307.
29. M. Aldape-Pérez, C. Yáñez-Márquez, O. Camacho-Nieto and A. J. Argüelles-Cruz, A New Tool for Engineering Education: Hepatitis Diagnosis using Associative Memories, *International Journal of Engineering Education*, **28**(6), 2012, pp. 1399–1405.
30. A.V. Uriarte-Arcia, I. López-Yáñez and C. Yáñez-Márquez, One-Hot Vector Hybrid Associative Classifier for Medical Data Classification, *PLoS ONE*, **9**(4), 2014, e95715.
31. C. López-Martin, I. López-Yáñez and C. Yáñez-Márquez, Application of Gamma Classifier to Development Effort Prediction of Software Projects, *Applied Mathematics & Information Sciences*, **6**(3), 2012, pp. 411–418.
32. B. Shahzad and A. M. Said, Identification and Quantitative Analysis of Project Success Factors for Large Scale Projects, *International Journal of Knowledge Society Research*, **5**(1), 2014, pp. 83–95.
33. C. Yáñez-Márquez, I. López-Yáñez, O. Camacho-Nieto and J. A. Argüelles-Cruz, BDD-based algorithm for the minimum spanning tree in wireless ad hoc network routing, *IEEE Latin America Transactions*, **11**(1), 2013, pp. 600–601.
34. I. López-Yáñez, O. Camacho-Nieto, C. Yáñez-Márquez and A. Uriarte-Arcia, Fast route convergence in dynamic power controlled routing for wireless ad–hoc networks, *IEEE Latin America Transactions*, **11**(1), 2013, pp. 607–608.
35. I. López-Yáñez, L. Sheremetov and C. Yáñez-Márquez, A Novel Associative Model for Time Series Data Mining, *Pattern Recognition Letters*, **41**, 2014, pp. 23–33.
36. C. Yáñez-Márquez, M. Aldape-Pérez, I. López-Yáñez and O. Camacho-Nieto, Emerging Computational Tools: Impact on Engineering Education and Computer Science Learning, *International Journal of Engineering Education*, **30**(3), 2014, pp. 533–542.
37. R. M. Bottino, Framing Technology Enhanced Learning Environments: Some Perspectives and Notions, *International Journal of Knowledge Society Research*, **3**(1), 2012, pp. 67–81.
38. V. Benson, S. Morgan and H. Tennakoon, A Framework for Knowledge Management in Higher Education Using Social Networking, *International Journal of Knowledge Society Research*, **3**(2), 2012, pp. 44–54.
39. M. D. Lytras and E. Kurilovas, Editorial: Special issue on Information and communication technologies for human capital development, *Computers in Human Behavior*, **30**, 2014, p. 361.
40. D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature*, **393**, 1998, p. 440.
41. P. Symeonidis, N. Iakovidou, N. Mantas and Y. Manolopoulos, From biological to social networks: Link prediction based on multi-way spectral clustering, *Data & Knowledge Engineering*, **87**, 2013, pp. 226–242.
42. Q. Yu, C. Long, C. Lv, H. Shao, P. He and Z. Duan, Predicting Co-Author Relationship in Medical Co-Authorship Networks, *PloS ONE*, **9**(7), 2014, e101214.
43. K. Narang, K. Lerman and P. Kumaraguru, Network flows and the link prediction problem, *ACM Proceedings of the 7th Workshop on Social Network Mining and Analysis*, Chicago, IL, USA, August 11–14, 2013, p. 3.
44. M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices, *Phys. Rev. E*, **74**(3), 2006, pp. 036104.

Ángel Ferreira-Santiago received his Computer Systems Engineering degree from the School of Computing of the National Polytechnic Institute, Mexico in 2013 and his Master's Degree in Computer Science in 2015 from the Center for Computing Research (CIC) of the same institution. He is currently pursuing a doctoral degree at the CIC. His research interests include evolutionary computation, pattern recognition, ensemble classifiers and computational intelligence.

Cornelio Yáñez-Márquez obtained his Bachelor degree (1989) on Physics and Mathematics at National Polytechnics Institute (IPN) Physics and Mathematics Superior School. His MSc (1995) and PhD (2002) degrees were received at IPN Center for Computing Research (CIC). Currently a Researcher Professor, Titular C, at IPN CIC. He was granted the Lázaro Cárdenas Award by the President of the Republic. Member of the National Researchers System (SNI). Areas of interest: Associative Memories, Neural Networks, Mathematical Morphology, and Software Engineering.

Itzamá López-Yáñez received his Bachelor degree as Information Systems Engineer (2003) at Monterrey Institute of Technology and Superior Studies (ITESM), while the MSc (2007) and PhD (2011) degrees on Computer Sciences at National Polytechnics Institute (IPN) Center for Computing Research (CIC), both with mention of Honor. He was granted the Lázaro Cárdenas 2012 Award by the President of the Republic. Currently he is a Professor at IPN Center for Technological Design and Development in Computer Science (CIDETEC), and member of the IEEE Communications Society. Areas of interest: Associative Memories, Neural Networks, Software Engineering, and Pattern Classification, in particular the Gamma classifier.

Oscar Camacho Nieto obtained his Bachelor degree as Communications and electronics Engineer (1989) at National Polytechnics Institute (IPN) Superior School of Mechanical and Electrical Engineering (ESIME), while the MSc degree (1995) on Computer engineering and the PhD degree (2003) on Computer Sciences, both at IPN Center for Computing

Research (CIC). Currently a Researcher Professor, Titular C, at IPN, and Director of the IPN Center for Technological Design and Development in Computer Science (CIDETEC). Areas of interest: Computer Architectures, Associative memories, Microprocessors, Digital Systems, and Neural Networks.

Mario Aldape-Pérez received his B.E. degree in electrical engineering from the Autonomous Metropolitan University, Azcapotzalco, Mexico, in 2004, and the M.Sc. degree in computer engineering and the Ph.D. degree in computer science from the National Polytechnic Institute of Mexico, Mexico, in 2007 and 2011, respectively. In 2010, he joined the Department of Computer Systems Engineering from the School of Computing, National Polytechnic Institute of Mexico, as an Associate Professor, and in 2012 became Professor. His current research interests include Unconventional Computing, Associative Memories, Soft Computing and FPGA Implementation of High Performance Pattern Classification algorithms. Dr. Aldape-Pérez is a member of the IEEE Computer Society. He is currently the President of the Academy of Digital Systems at the School of Computing of the National Polytechnic Institute of Mexico.

Amadeo-José Argüelles-Cruz received the M.Sc. degree in computer engineering and the Ph.D. degree in computer science from the National Polytechnic Institute of Mexico, Mexico, in 1997 and 2007, respectively. Since then, he has been with the Center for Computing Research (CIC-IPN), Mexico, where he is currently the Head of the Department of Research in Computer Science.