Communications of the IIMA

2012

Undergraduate Courses in Network Communications and Web Information Retrieval

Song Xing
Department of Information Systems California State University, Los Angeles

Xiannong Meng
Department of Computer Science Bucknell University

Wang Wei
School of Computer Science and Engineering Southeast University

Follow this and additional works at: http://scholarworks.lib.csusb.edu/ciima

Recommended Citation

Undergraduate Courses in Network Communications and Web Information Retrieval

Song Xing
Department of Information Systems
California State University, Los Angeles
sxing@calstatela.edu

Xiannong Meng
Department of Computer Science
Bucknell University
xmeng@bucknell.edu

Wang Wei
School of Computer Science and Engineering
Southeast University
Nanjing, Jiangsu, China
ww11@seu.edu.cn

ABSTRACT

In this paper we provide the result of an investigation on the teaching of computer networks courses and web information retrieval courses in undergraduate IT curriculum. The paper consists of two parts. First part of the paper surveys the undergraduate courses in two curriculum areas, computer networks and web information retrieval. The content of the survey includes the topics covered, the textbooks used, and major projects in these courses. The second part of the paper describes the authors' experiences in teaching such courses to undergraduate students in various discipline areas including computer science majors, IT majors, and non-technical majors.

Keywords: computer networks courses, web information retrieval courses, undergraduate IT curriculum, survey

INTRODUCTION

With the pervasive use of social networks and search engines, the knowledge of computer networks and web information retrieval are becoming ever more important in an undergraduate IT (Computer Science, Information Systems, etc) curriculum. Over the last 20 years since the World Wide Web was invented, the landscape of undergraduate computer networks courses and web information retrieval courses has been going through tremendous changes. It has been 10 years since the first major computer networks curriculum workshop (Kurose et.al., 2002). It also has been three years since a major survey in the teaching of web information retrieval courses (Fernandez-Luna et.al., 2009). In this paper we provide the result of an investigation on the teaching of computer networks courses and web information retrieval courses in undergraduate IT curriculum. The paper consists of two major parts. First part of the paper surveys the

undergraduate courses in two curriculum areas, computer networks and web information retrieval. The content of the survey includes the topics covered, the textbooks used, and major projects in these courses. The second part of the paper describes the authors' experiences in teaching such courses to undergraduate students in various discipline areas including computer science majors, IT majors, and non-technical majors. The rest of the paper is organized as follows. The course survey part consists of three sections. Section II gives an overview and a brief history of computer networks as well web information retrieval courses in undergraduate curriculum. Next, the methods of our study are described in Section III. Section IV presents the findings of our study that includes the textbooks used, the topics covered, and the approaches to cover these topics, the lab exercises, and the programming projects. The second part of the paper discusses the authors' experiences in teaching the computer network courses and web information retrieval courses to different undergraduate audiences. Section V presents one author's experiences in teaching a special topic course "Wireless Communications and Networks" to non-major students. Section VI and VII discuss second author's experiences in teaching web information retrieval courses to major students as a technical elective course and to first-year non-major students as a foundation seminar. We present our thoughts about teaching computer networks course and web information retrieval courses to undergraduate students in Section VIII.

A BRIEF OVERVIEW OF THE TEACHING OF COMPUTER NETWORKING COURSES AND WEB INFORMATION RETRIEVAL COURSES

This section presents the results of a study of sample computer network courses and web information retrieval courses offered in the last a few years in undergraduate curriculum. For the computer networks courses, the authors collected the information on the web from six universities that are representatives from either major research universities where the computer networks courses are available to the undergraduate students, or predominately undergraduate universities that offer computer networks courses to their students as an elective course. For the web information retrieval courses, the authors collected 40 courses from the web in the U.S.A. and in China and present a summary in the paper.

A. Computer networking courses

In 2002, SIGCOMM held the first ever workshop focused on the subject of computer networks education, entitled "Computer Networking: Curriculum Designs and Educational Challenges." (Kurose et.al., 2002). Eighty nine participants took part in the three panels of discussions, undergraduate curricula, laboratory-based courses, and graduate curricula. The workshop reports summarizes in a table a list of topics that a computer networking course should cover, as a minimum. These topics include physical network basics such as digital channels, coding, CSMA, and wireless networks; packet, circuit switching and framing; inter-networking; protocols; client-server programming; security and elementary performance measurement.

Ten years later, another SIGCOMM workshop on the same subject was held in August 2011 (ACM, 2011). The workshop includes four sessions that discuss various aspects of network education that touched a wide spectrum of topics. The workshop discussed some curriculum

topics such as using simulation tools to investigate large scale networks, introducing *network science* as a cross-discipline study, and teaching security along side with computer networks, among other topics. At the end of the workshop, the participants discussed where the computer network knowledge unit is placed in the overall computer science curriculum. This is partially in response to the call for computer science curriculum 2013 (Sahami et.al., 2012).

Net-centric knowledge units have been in the computer science curriculum since 2001 (ACM, 2001) and has been an essential part of a modern computer science curriculum. The subject of computer networking is taught from many perspectives. The one we most familiar with is from the computer science (or electrical engineering, information systems) point of view, which discusses the technical aspects of computer network, how the network is built, how the information is packaged and framed, how the packets travel from one place to another, how an application program can be written to accomplish certain functionality. However, computer networks can be studied from other perspectives. For example, how a business can use computer networks to enhance its operation; how today's popular media utilize computer networks to achieve their goals; how ordinary people can use social networks which operate on top of computer networks. The subject of computer networks used to be studied only by the computer science, computer engineering, electrical engineering, and information systems students. Since the proliferation of World Wide Web, especially, the social networks, students from all academic areas would like to learn something about computer networks. The expansion of the scope of teaching of computer networks presents computer networks educators with a tremendous opportunity, as well as challenges.

B. *Web information retrieval courses*

We review in this section the literature concerning the teaching and learning of web information retrieval. With increased interest in and importance of information retrieval and web search, more and more research projects have been on the subject of teaching and learning of information retrieval. Fenandez-Luna *et al.* (2009) presented a comprehensive review of the state of teaching and learning of information retrieval. In their paper, the authors presented taxonomy, educational goals, teaching and learning methods, assessment, and curricula regarding the subject of teaching and learning of information retrieval. The authors collected and analyzed 159 papers during the 40-year period of 1968 to 2008 that has anything to do with teaching and learning of information retrieval. The authors found that about 85 percent of these papers are from the field of library science and computer science. This collection of papers gives a trend in the evolution of information retrieval. While the authors didn't give a full list of the 159 papers, they did list 104 references in their survey paper, which is invaluable to the community of information retrieval education. The British Computer Society (BCS) held two international workshops on the teaching and learning of information retrieval (BCS, 2007; BCS, 2008). The proceedings of these two workshops contain rich collection of papers on various subjects related to the teaching and learning of information retrieval. In the 2007 workshop, a total of 10 papers were presented, topics ranging from learning environment (e.g., E-learning), to teaching strategies (math, IR, and web search), and to curricula and evaluation. In the 2008 workshop, a total of four papers were presented. The four papers discussed the topics of teaching IR as a philosophy problem, relation between search and engines, a holistic approach to teaching IR, and a report of developing a search engine as a practical project in teaching IR. In addition, other

papers have been published on the subject of teaching IR in recent years. McCown (2010) contrasted his experience in teaching an IR course in which students develop a search engine from scratch with the one in which students revise code in an existing search engine. Each of the two approaches has its pros and cons, developing a search engine from scratch gives students a greater understanding of what is behind the scene in a search engine but the end-product may be less polished, while revising existing search engine code may accomplish more functionality but students would have to overcome some steep learning curve. Zhu and Tang (2006) proposed a module-based integration of IR topics into different courses in an undergraduate curriculum. Meng (2003, 2011) presented the two cases teaching IR, one for computer science students who developed a search engine from scratch, the other for non-technical students who learned how to work with search engines and the societal impact as a result of the pervasive use of search engines.

Judging by the amount of publications and the number of courses available on the web, one can tell that overall interest in teaching and learning information retrieval in colleges and universities has been on the rise. While many aspects of teaching and learning information retrieval have been discussed in previous papers, we believe our unique contribution in this paper is to provide a survey of course contents, goals, textbooks used, and projects that are available on the web. Instructors who are interested in teaching such a course will find this collection of information useful in helping the development of a new course or revision of an existing course. Students who wish to learn the content on their own can also benefit from this collection of information.

METHODS OF OUR STUDY

For this study, a number of schools and course websites were first identified. We then chose courses whose websites are open to the public (some schools have their teaching materials behind a security gate, e.g., Moodle or Blackboard), from which we cannot access course materials. The authors manually studied the course teaching materials, collecting information such as syllabus, schedules, textbook and reference books, programming assignments, and hands-on laboratory exercises, from which a summary is presented in this paper.

A. Computer networking courses

We first concentrate on the computer networking courses. A large number of schools offer computer networking courses to their undergraduate students. While a comprehensive survey is of interest, we would like to find out what some representative schools are doing. We selected six schools as our samples.

We chose Princeton University and University of Massachusetts at Amherst because both universities have faculty members who published at least one popular textbook on the subject. We added Stanford University to this list as Stanford's computer networks course has been a source of information to the authors for a while and Stanford is known for its quality undergraduate education. Though many other universities fit the description mentioned above, they are not included in this study due to the limit in space and time.

We also chose a few schools which are primarily undergraduate institutions that offer strong computer science programs, most of which have engineering programs. (Bucknell University's computer science department offers computer science degrees in both of the College of Engineering and the College of Arts and Sciences. Administratively the department is a part of the engineering college.) These selected schools include Rose Hulman Institute of Technology, Richmond University, and Harvey Mudd College.

B. Web information retrieval courses

For the web information retrieval courses, the method of study is similar. The key difference is that while many undergraduate institutions teach computer networking courses, relatively few teach a course in web information retrieval to the undergraduate students. The authors were able to perform almost an exhaustive search for these courses on the web. The authors searched through the web for courses on the subject of information retrieval, web search, and web data mining. Each of these sites were manually visited, a few sites that didn't contain any technical content (e.g., websites that only listed a course title without any further information available on the web) were removed. We kept the sites in this survey that at the least we could identify the course title, instructor(s) of the course, and a list of main topics of the course. Most of the sites contain much richer contents than the aforementioned minimum amount of information. Among the additional information found on these sites include teaching schedules, topics discussed in the course, lecture notes, detailed homework and project assignments, and any combinations of the above. In the end, information from a total of 38 course websites is collected. These courses are mostly offered by U.S. universities with a few in Chinese universities.

STUDY RESULTS

We looked at a number of factors that may affect the curriculum in a typical undergraduate computer networks courses and web information retrieval courses. While there are no doubts that many factors would affect the outcome of an undergraduate course, we concentrated on the issues of textbooks and reference books, list of topics covered and the approaches to cover these topics, the lab exercises, and the programming projects.

A. Textbooks and reference books used in "Computer Networks"

A complete list of textbooks and reference books used by the six schools in their computer networks courses can be found in Appendix A. Two textbooks, one by Kurose and Ross (Kurose & Ross, 2009, 2012), and the other by Peterson and Davie (Peterson & Davie, 2007, 2011), seem to be the most popular ones. The text by Kurose and Ross is used by Rose Hulman, Richmond, Stanford, and UMass as the main text, while Harvey Mudd and Princeton use the text by Peterson and Davie. These two textbooks represent two difference approaches in teaching and understanding of computer networks.

The text by Kurose & Ross takes a top-down approach. After an introduction to the subject, the authors started the book with a chapter on the application layer using examples from the Internet and the Web. The application protocols studied at this level include HTTP, FTP, SMTP, and DNS. The chapter also investigates a popular application *Skype* as a P2P case study. A

discussion of general socket programming with which most application programs are created is presented towards the end of the chapter. The next subject presented by the text is the transport layer where the end-to-end protocols of TCP and UDP are discussed. Then what follows is the network layer where routing and addressing are discussed. The key protocol at this layer is the Internet protocol or IP. Common routing algorithms such as Link-State routing and Distance-Vector routing, as well as general hierarchical routing are presented. The lowest layer that is close to the access media is discussed in the next chapter, where the topics include coding, error detection and correction, media access control protocols, and address resolution. The book also contains a separate chapter on wireless and mobile network protocols, as well as a chapter on multimedia network. The top-down approach (from applications to media access) presents two major advantages. The first advantage is that the study starts with the applications such as the Web and Skype with which most students are familiar in their life before coming to a computer networks course, which motivates the students. The second advantage is that the top-down approach allows the course to study computer networks one layer at a time, from the most familiar one to the least familiar one and with the natural connections between higher level layers to the lower level layers.

The text by Peterson and Davie, on the other hand, uses a bottom-up approach in their book. The authors emphasize the building of blocks of lower layer protocols as the foundation of any higher layer protocols including such layers as applications. The authors using a bottom-up approach offer the readers a rigorous treatment of the computer networks from an engineering point of view. After a general introduction chapter to the topics, the book discusses the media access layer first, which includes a section on various wireless protocols. The discussion then moves to the issue of switching, followed by internetworking and transport layer protocols. The application layer (end-to-end data) is treated after the readers have an in-depth understanding how the networks function.

In addition to the two main textbooks discussed above, quite a number of different reference books are used by these courses. These books can be divided into three general categories, those that discuss computer networks in general (similar but in addition to the main text), those that help C programming in general, and those that help network programming in C specifically. Two other books are on the subject of general computer networking in addition to the two main textbooks. The book dealing with computer networks in general listed as a reference at Princeton is the book by Tanenbaum (2003). Tanenbaum's text is very well written, easy to read, and entertaining. It had been a very popular textbook before the two current texts by Kurose & Ross and Peterson & Davie became popular. Tanenbuam's book is very similar in organization to that of Peterson and Davie, that is, it takes a bottom-up approach, and the chapter organization is also very similar to that of Peterson and Davie. The popular C reference book by Kernighan and Ritchie (Kernighan & Ritchie, 1988*)* is used by Rose Hulman. This book is a classic reference book for the C programming language. If student programming projects and labs are required or suggested using C, this is an excellent reference text. Richmond uses Oualline's book (Oualline, 1997*)* as its C reference text. The rest of the reference books fall into the third category that guide students in network programming (Donahoo & Calvert, 2000; Stevens et.al., 2003; Stevens, 1994; Comer, 2000, 2003). These reference books mostly concentrate how to program network protocols using the programming language C.

B. Textbooks and reference books used in "Web Information Retrieval"

Throughout the web information retrieval courses we surveyed (see Appendix B for a compete list of textbooks in web IR), the textbook by Manning, Raghavan, and Schütze (2008) is by far the most popular one, 17 of the 28 courses use it as one of the main textbooks, including three universities in China. The other two popular books are the one by Baeza-Yates and Ribeiro-Neto (1999) (eight of 28) and the one by Croft, Metzler, and Strohman (2009) (five of 28). Both MRS and BYRN concentrate on the topics in information retrieval in general. MRS presents a more recent treatment of the topics than those of the BYRN as it is dated in 1999. Though the authors of BYRN have a new version of their book in 2011, the courses in our survey all quoted the book in its 1999 version at the time of our survey.

The authors of MRS aim the book at introductory level of graduate and upper level undergraduate students. The book contains a total of 21 chapters, each of which, according to the authors, can be covered in about one lecture unit of 75 to 90 minutes. The first eight chapters cover the core of information retrieval which includes retrieval models, index construction, term weights, ranking computation, and evaluation of retrieval. The second part of the book deals with more advanced topics using the foundation built in the first eight chapters. Various topics are discussed in this part, such as query processing, language models, classification and clustering, matrix decomposition, link analysis and other web search engine basics. BYRN, like MRS, starts with chapters that cover the basics of IR. The topics discussed in BYRN that are not in MRS include parallel and distributed IR, user interface and visualization, multimedia IR, and digital libraries. The book by CMS puts more emphasis on the science and engineering behind the application of the information retrieval in web search engines. The book uses web search engines as a vehicle to discuss the topics in IR. The book studies more algorithms and data structures related to information retrieval that are used in search engines. In addition to these three popular books, about 25 other books are used as main text or main reference in the surveyed the courses.

C. The topics and the covering approaches in "Computer Networks"

Although the main texts used in each of the courses contain a set of topics, not all topics are covered in the courses that use the textbook and the order of covering these topics varies. The course at University of Massachusetts at Amherst (UMass, 2011) which uses Kurose and Ross as its main text follows the order of the textbook closely, going from higher level protocols to lower ones. The courses at Rose Hulaman Institute of Technology (Rose Hulman, 2012) and Richmond University (Richmond, 2012) take a similar approach. In contrast to the top-down approach, the course at Harvey Mudd College (Harvey Mudd, 2010), while using Peterson and Davie as its main text, concentrates on the core of the network technologies. The course spends most time in lower level protocols from media access including wireless protocols, to switching, ATM, to IP, ARP, DHCP, ICMP, routing, DNS, to transport layer protocols (TCP, UDP). The course then spent three weeks on the counter part of these protocols in IPv6. The last part of the course (about 2.5 weeks) was dedicated to student team presentations. The course at Princeton University (Princeton, 2012) in general also takes a bottom-up approach as Harvey Mudd. However, the course swings from bottom layer protocols to the top ones from time to time (at least in the spring 2012 offering of the course). For example, the course discusses the topics of

HTTP protocol before going into the topics of routing, address resolution, and congestion control.

It is noteworthy that several courses keep their classes engaged and up-to-date on the latest technologies in computer networks. For example, the course at Stanford University (Stanford, 2011) follows the top-down approach using Kurose and Ross as its main text. However, the Stanford course covers some recent subjects not seen in other courses such as Data Center Networking. Large scale data centers are now an essential piece of infra-structure due to the huge amount of data available in many different settings, from search engines, to social networks, to e-commerce. The discussion of data center specific protocols brings students to the most up-to-date applications of network technologies.

D. The topics and the covering approaches in "Web Information Retrieval"

Since a course in the area of information retrieval and web search typically is an elective one, there are no required core components to cover, as one might find in other courses where the core is designated by the ACM and IEEE curriculum guidelines (ACM 2008). The exact topics vary from course to course, depending on the audience, the interests and expertise of the instructor(s), and other factors. Here we summarize the course topics in two groups, one focuses on the area of information retrieval, and the second focuses on search engines and related web technologies.

1. *Main topics of information retrieval:* Typical topics include text indexing, common retrieval models such as Boolean, vector, and probabilistic models, retrieval evaluation, query languages and operations, user modeling, and interface issues.
2. *Main topics of web search engines and technologies:* Typical topics include web search, crawling and indexes, link analysis, web meta-data, search engine architectures, web usage mining, spam and advertising, and social networks.

Most of the courses discuss a combination of the topics in the two main areas, information retrieval theory and its applications in software systems such as web search engines. A few courses are notably more tailored towards general information retrieval such as the one at CMU (Callan and Yang, 2011) and the one at UMass (Allan, 2010), while a few others are more explicitly on the subjects of web search such as the one at NYU (Davis, 2007) and the one at Harding University (McCown, 2009).

E. Lab exercises and programming projects in "Computer Networks"

Computer networks courses that we surveyed all contain a programming component, either through lab exercises, or through programming projects, or in some cases through both. In this sub-section, we look at the type of programming exercises or projects that each course requires.

At Rose Hulman, the programming component includes both lab exercises and a final team project. The topics in the labs cover various client-server programming at the socket level as well as error detection and correction exercises through socket programming. The last lab exercise

asks students to write a file transfer server/client program. The final project requires student teams to implement a *ping* utility and software routing forwarding function.

At Harvey Mudd, the lab exercises can roughly be divided into three groups, those that use Emulab (Emulab, 2012) as a tool to experiment with emulated networks; those that use Wireshark (Wireshark, 2012) to intercept and observe real network traffic; and those that use socket to write various applications.

The course at Richmond appears not requiring formal lab exercises. Instead, several projects are required. For example, the first project asks students to write a client-server program using socket. The second project involves writing a client-server program that multiple users can share their GPS locations. Students download and install an Android SDK on their computer in the third project and develop using the SDK a simple client-server program that sends a message to a server and receives the echo back from the server using the first project they developed.

The five lab exercises in Stanford's course are divided as follows. In the first two labs each asks students to implement a reliable transport protocol. The first one implements a stop-and-wait protocol, while the second one does a sliding window protocol. Lab 3 and 4 deal with routing, using a static routing table and dynamic routing, respectively. Lab 5 asks students to implement Network Address Translation (NAT).

At Princeton, there are also five programming assignments which are required. Differently, the first one is an introductory exercise of socket programming; the second assignment asks students to implement concurrent HTTP proxy server. Students implement an Internet router in the third project and a transmission control protocol in their fourth project. The last project asks students to measure the backbone network traffic.

Moreover, the course at UMass Amherst requires students to use Wireshark to monitor network for various types of traffic, HTTP, TCP, and UDP, etc. In addition, UMass requires students to complete two programming projects, one is a client-server application where the server has to be multi-threaded and the other asks students to implement distributed asynchronous distance vector routing.

F. Lab exercises and programming projects in "Web Information Retrieval"

In the case of web IR courses we found that the projects in the courses surveyed, when the descriptions are available, can be roughly divided into three categories, the ones that build a complete search system (simple or complex), the ones that modify a part or parts of an existing search system, and the ones that create a piece of software that functions as a stand-alone program to process, rank, or do other work on a body of text, but otherwise not as a complete search system.

Projects that fall into the first category, building a complete search system using a high level programming language include Davis (2007) in which students build a question-answer system using web content; Mihalcea (2011) in which students build a search engine within the UNT domain; and Yarowsky (2011) in which students can choose to build systems to find friends, to

classify news articles, or to help shopping. Projects that fall into the second category, modifying or creating a component to work with an existing search system include Agichtein (2010) in which students implement a ranking function for the Lucene open-source search engine; McCown (2009) in which students revise parts of the existing search engine Nutch; Strzalkowski (2011) in which students extend a selected component in Lucene such as term weight, page scoring, query expansion, and relevance feedback. Projects that fall into the third category include Callan and Yang (2011) in which students can choose to build personalized PageRank component, or a text classifier using Naïve Bayes method; Allan (2010) in which students are asked to write software to process, classify, index, and search a dataset from Enron employees' emails, the size of which is about half a million; Wilson (2008) in which students build an indexer through a series of smaller programming exercises.

SPECIAL TOPIC ELECTIVE AT CALIFORNIA STATE UNIVERSITY AT LOS ANGELS: WIRELESS COMMUNICATIONS AND NETWORKS

In this section, we describe the course and experience of the first author who taught a "Wireless Communications and Networks" course to non-majors.

Wireless technology has quickly become the newest networking technology that has hit the mainstream of communications systems. The purpose of this course is to provide a straightforward and broad survey of wireless voice and data network standards and technologies available today for personal and business communications. It is designed for business and information management students taking an entry-level wireless technology course or seeking better knowledge of wireless communications and networks, assuming the students having little or even no technical background.

The course is a major elective in Management of Information System at the business school and its prerequisite is *CIS 100 Business Computer Systems*, which is offered to all majors at the College of Business and Economics at California State University at Los Angles. CIS 100 provides students with computer system fundamentals, computer hardware and software concepts, and introduction to microcomputer software. Unlike other courses on wireless telecommunications systems and networks which provide a deeper understanding of the *operations of wireless technologies* used by professionals and technicians involved in a technical support area of mobile computing and wireless networking, this elective course emphasizes on the understanding of the *wireless networking systems concepts and principles*. The course is designed based on the assumption that the students taking this course have a basic knowledge of networking most likely from their non-technical experience using a wired or wireless computer network from Bluetooth and Wi-Fi to 3G/4G and satellite broadband at home or their business environment.

For the student learning objectives presented above, this course provides a fundamental introduction to all wireless communication systems including wireless personal area networks (infrared, Bluetooth, cordless phone), wireless local area networks (Wi-Fi), wireless metropolitan area networks (fixed and mobile WiMAX), and wireless wide area networks (cellular wireless, satellite communications). The topics cover the wireless networking systems'

architecture, standards, technologies, QoS (quality of service), security, and multimedia applications in business.

The course starts at the data and computer network communications overview, followed by a brief introduction to the data transmission techniques like RF (radio frequency) communications, signaling, modulation, and multiplexing. Although the course avoids the requirement of students for math, programming, and deep analysis of technologies, we believe the brief coverage of data transmission fundamentals will help to illustrate how they fit together in a modern wireless network. The course then majorly focuses on each wireless communication system listed above and ends at the wireless and personal communications applications in business which includes advantages of wireless technology, challenges of using wireless technology, wireless Internet access using Wi-Fi, WiMax, and 3G/4G, building a wireless infrastructure, radio frequency identification (RFID), wireless applications in medical and healthcare, industrial and commercial wireless applications, multimedia in wireless: audiovisual telephony, videoconferencing, broadcasting audio & video, and others.

The hands-on exercises on wireless LAN (local area network) devices, configuration and installation, troubleshooting, and maintain small and medium-sized wireless networks are assigned to students regularly accompanying to their homework assignments to help students establish the direct experience using wireless technologies and help them to better understand and master the subject matter of the topics. Students can even use their home network facilities to fulfill the hands-on projects. In addition, a group project is required of students, which builds upon and complementing the material covered in class.

We have chosen the text books with the writing style for non-major undergraduate students and which were tried by non-technical writing language. Such references include *Fundamentals of Wireless Networking* by Ron Price, McGraw-Hill/Irwin (Price, 2006), *Business Data Networks and Telecommunications* by Raymond Panko, Prentice Hall (Panko, 2008), and *Wireless# Guide to Wireless Communications* by Mark Ciampa and Jorge Olenewa (Ciampa and Olenewa, 2006), Cengage Learning.

The course topics we selected are also of assistance to students who are looking to obtain the Wireless# (or CWTS, Certified Wireless Technology Specialist) entry-level certification and the Certified Wireless Network Administrator (CWNA) foundation level certification from Planet3 Wireless, the organization that is the leader in vendor-neutral wireless certifications.

The College of Business and Economics at California State University at Los Angles has offered this course five times as a special topic since Spring 2007. The course has proved successful in student learning and attracted a large number of students each time from a variety of majors in the College of Business and Economics and other colleges in the university. The class was also very rewarding with high student evaluation scores each time.

TEACHING WEB INFORMATION RETRIEVAL TO MAJORS

This section discusses the experiences of the second author teaching web information retrieval course to computer science majors. This is a computer science elective course open to students who completed a junior level data structure and algorithm course. The textbook used was

Modern Information Retrieval by Baeza-Yates and Ribeiro-Neto (1999). A resource list is provided at the course website (Meng, 2006). The course contained 42 one-hour lecture periods. We presented most of the main topics in a typical information retrieval theory course. First we gave an overview of the information retrieval theory. Then we introduced one of the most exciting applications of the information retrieval theory, web search engines. Introducing an interesting application earlier motivates students to learn better the course materials. Students were able to see the connection between the general IR theory we discussed in the lectures and their actual applications. The basic vector space model was used to model the documents. We discussed indexing, retrieval evaluations, relevance feedback, web crawling, link structure analysis and general text properties. The key characteristic is that the lecture contents were closely matched with what were required for the particular phase of the course project. Towards the end of the course after students finish the programming project, students write a survey paper on various subjects of IR and web search, and present the findings to the class.

Student teams implement a simple, but functional search engine using a high level programming language. One important aspect of the project is to implement most of the functions using only the libraries that the chosen programming language supplies. Students are asked to parse the text using finite state machines, to build inverted index list using their own data structures, and to crawl the web using a depth-first or breadth-first graph traversal algorithm. The end product might not be a most efficient search engine. However students applied principles and techniques they learned in their previous computer science courses such as algorithms, data structures, and software engineering to complete the project.

TEACHING WEB INFORMATIONAL RETRIEVAL TO NON-MAJORS AS A FOUNDATION SEMINAR

In this section, we discuss the second author's experience in teaching the basics of web information retrieval to non-majors as a first-year foundation seminar.
Bucknell University requires its Arts and Science students to take a foundation seminar course during their first year (mostly in their first semester). Some Engineering students elect to take the seminar as well since it is very beneficial to the first-year engineering students.

While the general outcomes of a foundation seminar at Bucknell University are cultivating the capability of life-long learning through reading, writing, listening, presentation, and becoming information literate, the specific goals of this foundation seminar *FOUN 090-25: Search Engines and Our Lives* are to make students aware of the general technologies used in a typical web search engine, understand the advantages and limitations of using search engines, appreciate the societal implications brought in by the search engines. Through the exercise of reading, writing, presentation, and literature search in the seminar, students will become a better, more independent scholar after the seminar who is also knowledgeable about search engines.

We realize these outcomes by asking students to read papers, find extra references, synthesize what they read, write research papers, and present to the class of their findings. Because the seminar is intended for first-year non-major students in their first semester, the mathematics and computer science components from a typical web information retrieval course are removed from this seminar. Rather the seminar concentrates on general ideas of information retrieval and web

search engines. For example, instead of studying detailed algorithms and data structures for inverted index systems, an essential component of any search engine, we just illustrate the ideas of inverted indexing using diagrams and explain how they work in an information retrieval system. We also spent a half of the semester on societal impacts of search engines investigating human side of the issues related to web search such as politics, health care, privacy, environment, e-commerce, academics, among others.

We use the book *Search Engine Society* by Alexander Halavais (2009) as our main reference, accompanied by a number of other articles from research or popular publications. Alexander Halavais is a professor of communications at Quinnipiac University. His book contains a wealth collection of information about search engines and their social implications intended for general public. The book is suitable for first-year students of any major without background knowledge in computer science or information retrieval. The topics discussed in the seminar mainly follow the order of the book. We started with the basic building blocks of a search engine, followed by a brief history of search engines. We then presented a couple of examples of search engines, namely, AltaVista and Google, using research papers available to the public. The seminar then discussed various topics of search engine impacts on our lives listed at the end of previous paragraph. Instead of programming, students read, write, and present papers as they progress through the semester. Each week, students are asked to write a weekly reading journal based on the reading materials which can be a chapter from Halavairs' book, a conference paper, or a journal paper. Some chapters of the book edited by Amanda Spink and Michael Zimmer (2008) are also used as reading materials. Students then are asked to write and present two papers of their own, one in the subject of search engine technologies, the other in the subject of societal impact of these technologies.

One of the goals of the seminar is to introduce the basic working principles of a typical search engine to non-majors without involving math and programming. A search engine is abstracted as a working system with four major components, indexing, ranking, crawling, and user interface. When discussing indexing, the basic structure of inverted indexing system was presented so students understand for a given query, how a list of relevant documents can be generated. We also briefly discussed parsing of a text file. Students were asked to complete an assignment on paper to parse a given set of web pages into a collection of tokens and to build an inverted indexing system out of them. When discussing ranking, we talked about the elements that typically go into the ranking system such as page popularity using PageRank (Brin and Page, 1998), importance of the query words (location, frequency, and fonts of the query words in the document), and general ideas of term frequency and inverse document frequency. When discussing crawling, the subjects of traversing web pages using breadth-first, depth-first, or priority were studied. Students were asked to hand traverse sample web pages using breadth-first or depth-first algorithms. Also mentioned was the robot exclusion protocol so students understand that web crawlers are supposed to follow the protocol and respect web site owner's right. We studied a few robot.txt files on some popular websites such as www.cnn.com, www.abcnews.com, and www.ebay.com. When discussing user interface, we made a special point of client-server computing model where the browser is the client and the search engine is the server. We also discussed the fact that the server technically is able to, and many search engines do in fact, record user actions on the search engine. This discussion becomes a natural

transition to next segment of the seminar when we discuss user search behaviors and the issue of user privacy.

Throughout this segment of the seminar, we avoided programming and math details when search engine techniques were presented, concentrating only on the basic ideas. Students were able to grasp the ideas with written exercises such as traversing the web pages (graph), and computing PageRanks in its simple form (PageRank of a web page is the sum of PageRanks of other web pages pointing to it). Students were also able to construct correctly a basic inverted index system after applying parsing to a set of given web pages.

In the next segment of the seminar, we discussed search behaviors, that is, how users search the web and what kind of queries they have used in the past 10-15 years, using literature mostly from the research results of Spink and Jansen (Spink et.al. 2002, Spink and Zimmer, 2008, Jansen and Spink 2006). It was very natural for students to ask where the research data (user search queries) used in these research papers came from. The research data from Spink and Jansen's research came from search logs of a number of major search engines over the years. Students were surprised that search engines were able (and allowed) to log so much detailed information about web search and its users. We also discussed the incident in which AOL released search logs of about 658,000 users in July 2006 (Kawamoto and Mills, 2006) which raises serious privacy concerns for search users.

After gaining a basic understanding of how search engines work, the seminar turned its attention to the social impact brought in by the search engine technology. We examined the effect of search engines in areas such as e-commerce, politics as in presidential election and other political topics such as censorship in various countries, and cultural issues that different countries may have different views on the issue of freedom of speech on the internet, for example, environment issues where search engines use tremendous amount of energy and computer hardware in their data centers which result in environmental concerns, health care, and inequality in search among different segments of the society.

While a course project in a typical web search engine course for computer science majors would be a team programming project to build a simple search engine, the main student wok in this foundation seminar consists of various reading and writing assignments. The writing assignments helped students reflect on what they learn in the seminar. The assignments also are used to meet the writing goal for the foundation seminar. Two types of writing assignment are used in the seminar. One is the weekly reading journal. Students are asked to write a reading journal of 300 to 500 words every week on the subject of what was being discussed during that week. The other type of writing assignment is the research papers. Each student wrote two research papers during the semester, one written in team on the subject of search engine technology and its history which was due about half-way through the semester; the other written as an individual paper on the subject of social impact of search engines which was due at the end of the semester. The research papers are typically 1,200 to 2,500 words in length.

CONCLUSION

This paper presents a glimpse of the current state of computer network courses and web information retrieval courses in undergraduate IT related curriculum. The authors selected a number of different schools whose computer networks course materials and web information retrieval courses materials are available online. The information presented in this paper includes the textbooks and reference books, the main topics and how they are covered, and the laboratory and programming projects used in these courses. From the survey we found that the textbooks and the teaching approaches are relatively concentrated on two main ones, top-down or bottome-up, while the types of lab and programming assignments are in a wide variety. The motivation of the paper is to gather information from across the Internet about computer network courses and web information retrieval courses to see how currently the courses are taught in different schools and use the information collected as guidelines for course development and revision. The authors also share their own experiences in teaching these two courses to different audience in the hope that we can exchange ideas with colleagues in the field.

REFERENCES

ACM. (2001). *Computing Curricula 2001*. Retrieved from http://www.acm.org/education/curric_vols/cc2001.pdf

ACM (2008). *Computer Science Curriculum 2008* (http://www.acm.org//education/curricula/ComputerScience2008.pdf)

ACM. (2011). *SIGCOMM 2011 Education workshop*, Retrieved from http://edusigcomm.info.ucl.ac.be/Workshop2011/Workshop2011

Agichtein, E.(2010). *CS572: Information Retrieval and Web Search* (http://www.mathcs.emory.edu/~eugene/cs572/) at Emory University, Spring 2010.

Allan, J. (2010). *Information Retrieval* (http://cs646.cs.umass.edu/) at University of Massachusetts, Fall 2010.

Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.

BCS. (2007). *Proceedings of the First International Workshop on Teaching and Learning of Information Retrieval (TLIR 2007). Available at: http://ewic.bcs.org/category/16371*.

BCS. (2008). *Proceedings of the Second International Workshop on Teaching and Learning of Information Retrieval (TLIR 2008). Available at: http://ewic.bcs.org/category/16295*.

Brin, S. & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. In *Proceedings of the 7th IWWW Conference*, Brisbane, Australia, 14-18 April 1998.

Callan, J. and Yang, Y. (2011). *11-741: Information Retrieval* (http://boston.lti.cs.cmu.edu/classes/11-741/index.html) at Carnegie Melon University, Spring 2011.

Ciampa, M. and Olenewa, J. (2006). *Wireless# Guide to Wireless Communications*. Cengage Learning.

Comer, D.E. (2000). *Internetworking with TCP/IP Vol 1: Principles, Protocols, and Architecture,* Fourth Edition, Prentice Hall.

Comer, D. E., (2003). *Hands-on Networking with Internet Technologies,* 2nd Edition, Prentice Hall.

Croft, W.B., Metzler, D. and Strohman, T. (2009). *Search Engines: Information Retrieval in Practice*. Addison Wesley.

Davis, E. (2007). *G22.2580: Web Search Engines* (http://cs.nyu.edu/courses/fall07/G22.2580-001/index.html) at New York University, Fall 2007.

Emulab. (2012). *Network Emulation Testbed*. Retrieved from http://www.emulab.net/

Fernandez-Luna, J.M., Huete, J.F., MacFarlane, A., and Efthimiadis, E.N. (2009). Teaching and learning in information retrieval. *Information Retrieval.* 12:201–226.

Halavais, A. (2009). *Search Engine Society* Malden, MA: Polity Press.

Harvey Mudd College, *CS 125: Computer Networking.* http://www.cs.hmc.edu/~mike/courses/cs125/f10/index.html

Jansen, B.J. & Spink, A. (2006). How Are We Searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs. *Information Processing & Management*, 42(1), 248-263.

Kawamoto, D. & Mills, E. (August 7, 2006). AOL apologizes for release of user search data. *CNET News.* Accessed December 29, 2009 from http://news.cnet.com/2100-1030_3-6102793.html

Kernighan B. W. & Ritchie, D.M. (1988). *C Programming Language, Second Edition*, Prentice Hall.

Kurose, J., Liebeherr, J., Ostermann ,S., & Ott-Boisseau, T. (2002). *ACM SIGCOMM Workshop on Computer Networking: Curriculum Designs and Educational Challenges.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.109.4558

Kurose, J.F. & Ross, K.W. (2009). *Computer Networking: A Top-Down Approach Featuring the Internet (fifth edition)*, Addison-Wesley.

Kurose, J.F. & Ross, K.W. (2012). *Computer Networking: A Top-Down Approach Featuring the Internet (sixth edition)*, Addison-Wesley.

Manning, C., Raghavan, P., and Schutze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

McCown, F. (2009). *COMP 475: Search Engine Development* (http://www.harding.edu/fmccown/classes/comp475-s09/) at Harding University, Spring 2009.

McCown, F. (2010). Teaching web information retrieval to undergraduates. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education* (SIGCSE 2010), Mar 2010, Milwaukee, WI, pp. 87-91. doi:10.1145/1734263.1734294

Meng, X. (2003). Putting information retrieval theory into practice—A web search engine project for an undergraduate computer science elective course. In *Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition.*

Meng,   X.   (2006).   *CSCI   335:   Web   Information   Retrieval* (http://www.eg.bucknell.edu/~csci335/2006-fall/index.html) at Bucknell University, Fall 2006.

Meng, X. and Xing, S. (2011). Teaching Web Information Retrieval and Network Communication Technology to Non-Major Undergraduate Students. In *Proceedings of the 2011 ASEE Annual Conference and Exposi*tion, Vancouver, B.C., Canada.

Mihalcea,   R.   (2011).   *CSCE   5200   Information   Retrieval   and   Web   Search* (http://www.cse.unt.edu/~rada/CSCE5200/) at University of North Texas, Spring 2011.

Oualline, S. (1997). *Practical C Programming* (3rd Edition), O'Reilly.

Panko, R. (2008). *Business Data Networks and Telecommunications*, Prentice Hall.

Peterson, L. & Davie, B. (2007) *Computer Networks, A System Approach, 4th Edition*, Morgan Kaufmann.

Peterson, L. & Davie, B. (2011) *Computer Networks, A System Approach, 5th Edition*, Morgan Kaufmann.

Price, R. (2006) *Fundamentals of Wireless Networking*, McGraw-Hill/Irwin.

Princeton   University.   (2012).   *COS   461:   Computer   Networks.* http://www.cs.princeton.edu/courses/archive/spring12/cos461/

Richmond   University.   (2012).   *CMCS   332:   Computer   Networks.* https://facultystaff.richmond.edu/~dszajda/classes/cs332/Spring_2012/

Rose Hulman Institute of Technology. (2012). *CSSE 432 – Computer Networks* http://www.rose-hulman.edu/class/csse/csse432/201230/

Sahami, M., Roach, S., Cuadros-Vargas, E., & Reed, D. (2012). *Computer science curriculum 2013: reviewing the strawman report from the ACM/IEEE-CS task force.* Retrieved from http://dl.acm.org/citation.cfm?id=2157140&dl=ACM

Spink, A., Jansen, B.J., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *Computer*, 35(3), 107 - 109.

Spink, A. & Zimmer, M. (2008). *Web Search -- Multidisciplinary Perspectives*. Berlin Heidelberg: Springer-Verlag.

Stanford   University.   (2011).   *CS144:   Introduction   to   Computer*   Networking. http://www.scs.stanford.edu/11au-cs144/notes/

Stevens, W. R., Fenner, B., & Rudoff, A.M. (2003). *UNIX Network Programming: The Sockets Networking API, 3rd Edition,* Addison Wesley.

Stevens, W.R. (1994). *TCP/IP Illustrated, Volume 1: The Protocols,* Addison-Wesley.

Strzalkowski, T. (2011). *CSI 550: Information Retrieval* (http://aquarius.ils.albany.edu/~minoo/csi550/) at University of Albany, Fall 2011.

Tanenbaum, A.S. (2003). *Computer Networks* (4th edition), Prentice Hall.

University of Massachusetts. (2011). *CS453: Computer Networking* http://www-net.cs.umass.edu/cs453_fall_2011/

Wilson, G.V. *Information Retrieval* (http://www9.georgetown.edu/faculty/wilsong/IR/IR.html) at Georgetown University, Spring 2008.

Wireshark. (2012). *Network Protocol Analyzer*. Retrieved from http://www.wireshark.org/

Yarowsky, D. (2011). *Information Retrieval and Web Agents* (http://www.cs.jhu.edu/~yarowsky/cs466.html) at Johns Hopkins University, Spring 2011.

Zhu, L., & Tang, C. (2006). A module-based integration of Information Retrieval into undergraduate curricula. *Journal of Computing Sciences in Colleges*, 22(2), 288–294.

APPENDIX A:

Textbooks and Reference Books for "Computer Networks"

Comer, D.E. (2000). *Internetworking with TCP/IP Vol 1: Principles, Protocols, and Architecture,* Fourth Edition, Prentice Hall.

Comer, D. E., (2003). *Hands-on Networking with Internet Technologies,* 2nd Edition, Prentice Hall.

Donahoo, M.J. & Calvert, K.L.(2000). *Pocket Guide to TCP/IP Socket Programming in C*, Morgan Kaufmann.

Donahoo M.J. & Calvert, *K.L. (2009). TCP/IP Sockets in C: Practical Guide for Programmers* (2nd Edition), Morgan Kaufmann.

Fall, K. R. & Stevens, W.R. (2011).*TCP/IP Illustrated, Volume 1: The Protocols* (2nd Edition), Addison-Wesley.

Kernighan B. W. & Ritchie, D.M. (1988). *C Programming Language, Second Edition*, Prentice Hall.

Kurose, J.F. & Ross, K.W. (2009). *Computer Networking: A Top-Down Approach Featuring the Internet (fifth edition)*, Addison-Wesley.

Kurose, J.F. & Ross, K.W. (2012). *Computer Networking: A Top-Down Approach Featuring the Internet (sixth edition)*, Addison-Wesley.

Oualline, S. (1997). *Practical C Programming* (3rd Edition), O'Reilly.

Peterson, L. & Davie, B. (2007) *Computer Networks, A System Approach, 4th Edition*, Morgan Kaufmann.

Peterson, L. & Davie, B. (2011) *Computer Networks, A System Approach, 5th Edition*, Morgan Kaufmann.

Stevens, W. R., Fenner, B., & Rudoff, A.M. (2003). *UNIX Network Programming: The Sockets Networking API, 3rd Edition,* Addison Wesley.

Stevens, W.R. (1994). *TCP/IP Illustrated, Volume 1: The Protocols,* Addison-Wesley.

Tanenbaum, A.S. (2003). *Computer Networks* (4th edition), Prentice Hall.

Appendix B

Textbooks and Reference Books for "Web Information Retrieval"

Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley .

Baldi, P., Frasconi, P., and Smyth, P. (2003). *Modelling the Internet and the Web*. John Wiley and Sons.

Battelle, J. (2005). *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. Portfolio Hardcover.

Belew, R. K. (2001). *Finding Out About: A Cognitive Perspective on Search Engine Technology and the WWW*. Cambridge University Press.

Bourne, C. P. and Hahn, T.B. (2003). *A History of Online Information Services: 1963-1976.* The MIT Press.

Buettcher, S., Clarke, C.L.A., and Cormack, G.V. (2010). *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press.

Chakrabarti, S. (2002). *Mining the Web: Discovering Knowledge from Hypertext Data*. Morgan Kaufmann.

Chang, G. (2001). *Mining the World Wide Web*. Springer.

Cheong, F. (1996). *Internet Agents: Spiders, Wanderers, Brokers, and Bots.* Indianapolis, IN : New Riders.

Croft, W.B., Metzler, D. and Strohman, T. (2009). *Search Engines: Information Retrieval in Practice*. Addison Wesley.

Croft, W. B. (ed). (2000). *Advances in Information Retrieval: Recent Research from the Center for Intelligent Information Retrieval.* Kluwer Academic Publishers.

Frakes, W. and Baeza-Yates, R. (1992). *Information Retrieval: Data Structures and Algorithms*. Englewood Cliffs, N.J. : Prentice Hall.

Greengrass, E. (2000). *Information Retrieval: A Survey.* Available online at: http://www.csee.umbc.edu/csee/research/cadip/readings/IR.report.120600.book.pdf.

Grossman, D. (n.d.). *Information Retrieval*. Available online at: http://ir.iit.edu/~dagr/cs529/ir_book.html

Han, J. and Kamber, M. (2000). *Data Mining - Concepts and Techniques*. Morgan Kaufmann.

Hearst, M. (2009). *Search User Interfaces*. Cambridge University Press. Available online at: http://searchuserinterfaces.com/

Hersh, W. R. (2003). *Information Retrieval: A Health and Biomedical Perspective*. 2nd Edition. Springer-Verlag.

Korfhage, R. R. (1997). *Information Storage and Retrieval.* John Wiley & Sons.

Kowalski, G. and Maybury, M. T. (2000). *Information Storage and Retrieval Systems: Theory and Implementation*. Kluwer Academic Publishers.

Langville, A. N. and Meyer, C.D (2006). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton Press.

Levene, M. (2005). *An Introduction to Search Engines and Web Navigation*. Pearson.

Liu, B. (2011). *Web Data Mining*. Springer.

Manning, C., Raghavan, P., and Schutze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Marchionini, G.(1997). *Information Seeking in Electronic Environments*. Cambridge University Press.

van Rijsbergen, C. J. (1979). *Information Retrieval*. Available online at: http://www.dcs.gla.ac.uk/Keith/Preface.html

Salton, G. (1988). *Automatic text processing: the transformation, analysis, and retrieval of information by computer.* Reading, Mass. :Addison-Wesley.

Sparck Jones, K. and Willett, P. (1997). *Readings in Information Retrieval.* Morgan Kaufmann.
van der Weide, T. (2001). *Information Discovery*. Available online book at: http://osiris.cs.kun.nl/iris/web-docs/edu/ir1/ir1.pdf

Witten, I. H., Moffat, A., and Bell, T.C. (1999). *Managing Gigabytes. Available online at:* http://ww2.cs.mu.oz.au/mg/

Wong, C. (1997). *Web Client Programming*. O'Reilly and Associates. Available online at: *http://oreilly.com/openbook/webclient/*