CS 4675 –Internet Computing Systems, Services, and Applications

Overview:
The purpose of this course is to give students a thorough introduction to the cutting edge technology used in designing and developing in Internet computing systems, services and applications. It is also expected that students will gain experience and practice with hand-on course projects and critiques to assigned reading of research papers. The focus in the course is on developing skills and experience in understanding, designing, and developing advanced Internet Computing Systems and applications.

Prerequisites:
Students are expected to have taken Operating Systems (CS2200 or equivalent) and Computer Networks. Parallel and Distributed Systems (CS 4230/6236) and High Performance Parallel Computing (CS 6230) are highly recommended for students interested in doing research in systems or networks. In addition, students are expected to have a solid grasp of Java/C/CGI programming. Sockets programming is not required but desirable.

Course Objectives
This course reviews concepts, techniques, and systems issues in advanced Internet systems, services and applications, and explores new challenges and research issues that are critical for designing and developing of basic and advanced Internet computing systems and services. Examples of basic Internet services include Internet scale search services (including search engines, crawling robots and indexing servers), Web servers, application servers (such as the E-commerce information servers like eBay, CNet, Amazon), and web-based online transaction systems. Examples of advanced Internet Systems include mobile and wireless systems and services, sensor network applications, pervasive computing applications and grid and peer to peer systems and applications. Main themes of the course include fundamentals of these basic and advanced Internet systems and applications, especially important search techniques, indexing techniques, various mechanisms for querying heterogeneous data sources and delivering fresh information from sources to consumers, and intelligent methods for distribution of control to avoid bottlenecks and enhance scalability, security and reliability of massively distributed Internet systems. One of the important goals of the course is to look beyond the present status of the Internet and conjecture what possible future technologies and applications will evolve. The course will include a significant project component that will typically require Java/C/Script programming.

Prerequisites:

Operating Systems	File system services, scheduling, and memory management
Computer Networks	Basics of TCP IP, Network protocols, and Network Management

Topical Coverage

Course Introduction and Evolution of Internet Computing Systems (1 lecture)

Fundamentals of Content Delivery and Dissemination on the Internet (1 week)
- Basic models of information delivery and dissemination: request-respond, publish-subscribe, broadcast.
- Techniques and optimizations commonly used in each of the three delivery models.

Web Servers and Application Servers Technology (2 weeks)
- Web Servers
- Web Proxy Servers
- Application Servers
- Web Server Clusters, including local and global web server clusters and scheduling algorithms

Search Engines Technology (1 week)
- Web Crawlers
- Index Engine
- Search Engine Ranking Algorithms, including query independent algorithms like PageRank and query dependent algorithms likes HITS.
- Meta Search Engines and Domain-specific search engines
- Search Engine Comparison Techniques

Peer to Peer Computing Systems and Applications (2 weeks)
- Centralized P2P – Napster like
- Decentralized unstructured P2P – Gnutella-like
- Decentralized structured P2P – Chord like
- Convergence of Grid Computing and P2P computing

Publish-Subscribe Systems and Applications (1 week)
- Centralized Publish-subscribe systems
- Distributed publish-subscribe systems
- Secure publish-subscribe systems

Mobile Computing Systems and Applications (2 weeks)
- Mobile Network Basics
- Mobile Location-based Services
  - Location data types, location models, location acquisition, and location service architectures
- Broadcast Services: Scheduling Algorithms and Energy-effcient Optimizations
- Mobile Peer to Peer Systems
- Location Privacy and Security

Sensor Network Computing Systems and Applications (1 week)
- Sensor Networks: The basics
- Sensor Data Acquisition Systems
- Sensor Network Applications
  - Environmental monitoring, Disaster recovery, Transportation, RFID
- Sensor Network Systems Security

Systems issues in Internet Computing (4 weeks)
- Middleware (including registration and discovery services)

- Load balancing and throughput maximization
- Dynamic adaptations, including parameterized optimization, code composition, code compilation and recompilation, runtime code injection)
- Dynamic resource management (CPU, memory, network bandwidth)
- Fault tolerance, system-level recovery and application-level recovery
- Reliable computing with unreliable components
- Internet Security and data privacy

Internet Measurements **(1 lecture)**
- Latency and throughput
- Performance
- Scalability
- Reliability
- Power management

Textbook

- None

Required readings
- Handouts made available to the class via course notes and published papers

Evaluation

- Exam   30%
- Homework   20%
- Projects   50%