Design Project
Stage 0

Overview

This document provides a brief outline for the course design project. Further documents will be released to provide you the problem description and to guide you through specific stages of the design.

The design project is to be performed by teams of four people each. Five teams will be formed. Team formation is left up to you. I recommend forming a team with an eye for completing both this project and the supplemental data analysis project, rather than having different teams on these projects.

The project will have the following timeline:

No.	Activity/Assignment	Dates/Due Date	Comment
1.	Initial discussion	*April 16, 2009*	brief discussion during the lab
2.	Team formation	*by April 21, 2009*	
3.	Problem description released	*April 21-23, 2009*	
4.	Stage 1 (preparation)	*April 23-28, 2009*	
5.	Discussion	*April 28, 2009*	*tentative*
6.	Stage 2 (design)	*April 30 - June 4, 2009*	
7.	Report due	*June 4, 2009*	
8.	Presentation	*June 10, 2009*	final exam slot

We may use a full lab or part of a lab session some time in May to conduct more discussions on the project.

Brief Outline of the Problem

This is a brief outline of the problem. The full specification will be released next week, together with the Stage 1 assignment.

A bright (?) idea. Those of you who have tried to make a connection in an overcroweded airport knows just how big a hassle it may be. There is a technological solution that, if implemented, can pave the way for significant reduction in airport hassle. The idea is simple: *embed an RFID chip into every boarding pass issued by airlines.* RFID chips are thus associated with specific legs of specific trips that passengers take.

RFID chips are trackable on the airport scale, and can provide information in real time about *whether* specific passengers are at the airport and *where* they are. The may allow for targeted information delivery to the passengers, and they may allow airlines to determine whether, for example, it makes sense to wait for a passenger or to close to doors of a departing plane.

Your mission, should you choose to accept it... Obviously, keeping track of RFID chips embedded into boarding passes falls onto the lap of software. Your goal is to eventually *design the architecture for a software system* that

- Enables the use of RFID-enhanced boarding passes across multiple airlines, airports, etc...
- Provides a wide range of management and control facilities that take advantage of the features provided by RFID chips.
- Collects, stores, maintains and **analyses** information that arises in the course of running the software.

And this is relevant because... Data Mining/Knowledge Discovery tasks rarely occur in vacuum. They are usually an integral part of a larger software system, or a conglomerate of software systems, whose primary role is operational in nature. Such large systems generate vast amounts of data in the course of their operation. This data is archived and used for analytical purposes.

The RFID enhancement of boarding passes is a realistic, but not yet realized scenario, which has the potential of generating vast quantities of data, which, in turn, can be analyzed for various purposes. Throughout the project you will have to determine

- What types of information may be available to you;
- How this information can be collected/stored/maintained/archived;
- What analytical questions you need/want answered;
- What types of intelligent data analysis/knowledge discovery/data mining tasks you can perform with the available data in order to answer those questions.

So, while the final deliverable of your project will be the design of the overall system, including **all the non-KDD bits**, it is expected that the significant amount of effort in your design will be devoted to resolving the issues of data storage, integration and analysis.

Stage 0 Assignment

1. **Organize into teams.** I am looking to have five teams of four people each. Team formation is left up to you. As soon as each team is created, please create a wiki page with the team information (names of team members, team name, etc.). **All teams must be organized by April 21.**

2. **Stakeholders.** While you will mostly be working on software design, it is useful to understand the role of the software in the broader context. I would like to ask each of you individually, or collectively, to think about various **stakeholders** in the process of developing and deploying the RFID-enhanced boarding pass technology. I would like to ask each team have a preliminary list of stakeholders by April 23. We will take a few mins during the April 23 lab period to discuss the Stage 1 assignment, and some aspects of that assignment will be predicated upon our joint understanding of all the stakeholders in this process.