

CIS400/401 Project Proposal Specification

Dept. of CIS - Senior Design 2009-2010

Insup Lee*
lee@cis.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

Andrew G. West
westand@cis.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

Alex Roederer
roederer@cis.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

ABSTRACT

*Abstract: The purpose of this document is three-fold. First, it describes the requirements of the project proposal. Second, it presents this information in a style your proposal should mimic. Third, by viewing the *.tex ‘code behind’, some basic L^AT_EX technique can be learned. Do not feel compelled to completely follow our style – but do use it as a guide.*

Like this ‘report’, your proposal is expected to have an abstract. An abstract is a two-paragraph maximum executive summary of your work. It should briefly outline your problem statement and its (expected) contributions.

1. INTRODUCTION

You should begin by introducing your topic. You should define core terminology specific to the field, introduce the problem statement, and make clear the benefits (motivate!) of resolving that problem statement.

Let us suppose you are attempting to build a constant time integer factorization algorithm. You would want to briefly explain that factorization is the process of breaking an integer into its prime multipliers. Give a simple example. Your problem statement would conclude that factorization is an unnecessarily complex process. Finally, you would state that via your proposed system, factorization will be easy and the consequences that will have. Namely, the RSA-encryption algorithm will be broken.

2. RELATED WORK

Perhaps the most important section of your proposal is *related work*. Here you demonstrate that you have read and understand what others in the field have done. This ensures you (1) know the state-of-the-art, (2) are not re-doing others work, and (3) you know the performance levels you must achieve to make a contribution. As you discuss each related work, make note of how each has advanced the field. More importantly, note their shortcomings – these you will exploit to come up with a *better* system.

This section should have in-line citations to your bibliography (really all sections should have citations, but we expect them to be most dense in this section). We are going to require that your proposal has at least 6 references. Fortunately, L^AT_EX makes citations easy. Your two TAs have had no difficulty, as the work of West *et al.* [2] demonstrates. Need help with L^AT_EX? Be sure to check out [3] and [1], two helpful on-line resources.

*List your advisor (last) amongst the authors. Your advisor must also sign a hard-copy of your proposal.

What defines a good resource? Wikipedia is **NOT** a good resource. We would like to see references from academic journals/conferences (ACM, IEEE, etc.). We realize not everyone is doing pure research and for students with ‘implementation’ projects such sources may be rare. No matter the case, your sources need to be reputable. You will later be asked to produce an *annotated bibliography* where you defend the quality of your resources.

Let us return to your factorization proposal. You should put out the earliest related work; naïve methods like trial division and the Sieve of Eratosthenes, but state they are of no modern relevance. Then discuss modern methods like the Quadratic Sieve¹ and General Number Field Sieve. Note the humongous time and memory bounds of these algorithms. But wait! You are going to propose a better way . . .

3. PROJECT PROPOSAL

Now is the time to introduce your proposed project in all of its glory. Admittedly, this is not the easiest since you probably have not done much actual research yet. Even so, setting and realizing realistic research goals is an important skill. Begin by summarizing what you are going to do and the expected benefit it will bring.

3.1 Anticipated Approach

Having summarized *what* you are going to do, its time to describe *how* you plan to do it. Our factorization example does not work so well here (its likely impossible to realize) – so let us suppose you are going to create a service that takes a cell-phone picture of a building and returns via text-message, the name of that building².

In this case you might want to talk about establishing a server to receive pictures via MMS. Once the picture is received, you will run an edge extraction algorithm over it. Then, similarity between the submitted picture and those stored (and tagged) in a MySQL database will be computing using algorithm *XYZ*. Finally, the tag of the most similar image will be returned to the user. Do not bore the reader with trivial details, but give them an overview; a block-flow diagram may prove extremely helpful.

Also note where you anticipate having novel difficulty. Maybe you have never setup a MySQL database or even used SQL before at all – yes, that is a challenge – but not one reader’s care about. More novel would be the fact that

¹This was the topic of one TAs undergraduate thesis, and the only reason this particular example was chosen.

²Do not use this idea – someone did it last year.

many buildings on Penn’s campus look similar and your classifier may be inaccurate in such instances.

3.2 Evaluation Criteria

Suppose you have implemented your approach, and it is functioning. Now how are you going to convince readers your approach is better than what exists? In the factorization example, you could just compare run-times between algorithms run on the same input. The image recognition example might use a percentage of accurate classifications. Other fields may have established testing benchmarks.

No matter the case, you need to prove you have contributed to the field. This will be easier for some than others. In particular, those with ‘sensory’ projects involving visual or sonic elements need to think this point through – objectivity measures are always better than subjective ones.

4. RESEARCH TIMELINE

Finally, we would like you to speculate about the pace of your research progress. This section need not be lengthy, we would just like you to specify some milestones so we can gauge your progress during our intermediate interviews. Let us follow through with our image recognition example:

- ALREADY COMPLETED: Preliminary reading. Began implementation of image-recognition algorithm.
- PRIOR-TO THANKSGIVING : Photograph buildings for DB. Make algorithm more efficient, tune parameters.
- PRIOR-TO CHRISTMAS : Create server-MMS interface. Expand tagged DB collection.
- COMPLETION TASKS : Verify implementation is bug-free. Conduct accuracy testing. Complete write-up.
- IF THERE’S TIME : Investigate image pre-processing techniques to improve accuracy.

5. REFERENCES

- [1] The Comprehensive TeX Archive Network (CTAN). A (not so) short introduction to LaTeX2e. <http://www.ctan.org/tex-archive/info/lshort/english/>.
- [2] Andrew G. West, Adam J. Aviv, Jian Chang, Vinayak S. Prabhu, Matt Blaze, Sampath Kannan, Insup Lee, Jonathan M. Smith, and Oleg Sokolsky. QuanTM: A quantitative trust management system. In *EUROSEC 2009*, pages 28–35, March 2009.
- [3] Wikibooks. LaTeX. <http://en.wikibooks.org/wiki/LaTeX>. Note: Students should not cite Wikis!

APPENDIX

A. OTHER SPECIFICS

Your proposal need not have appendices like this section and the next, but we still have critical info to share:

1. PROPOSAL LENGTH: We require that your proposal be 4–5 pages in length, bibliography included. Be careful, L^AT_EX and our style-file in particular are *extremely* space efficient. An 9-page MS-Word document could easily become a 5-page L^AT_EX one.
2. PLAGIARISM: **DO NOT** plagiarize. If you are caught, you will fail the class (*i.e.*, not graduate), or worse.

B. L^AT_EX EXAMPLES

At this point, the proposal specification is complete. From here on out, we are just going to show off some commonly used L^AT_EX technique. Be sure to look at the ‘code behind’ and see Tab. 1, Eqn. 1 and Fig. 1 for the output!

$$M(p) = \int_0^\infty (1 + \alpha x)^{-\gamma} x^{p-1} dx \quad (1)$$

| User Type | Cleanup% | Honesty% |
|---------------------|----------|------------|
| Good | 90-100% | 100% |
| Purely Malicious | 0-10% | 0% |
| Malicious Provider | 0-10% | 100% |
| Feedback Malicious | 90-100% | 0% |
| Disguised Malicious | 50-100% | 50-100% |
| Sybil Attacker | 0-10% | Irrelevant |

Table 1: Example Table

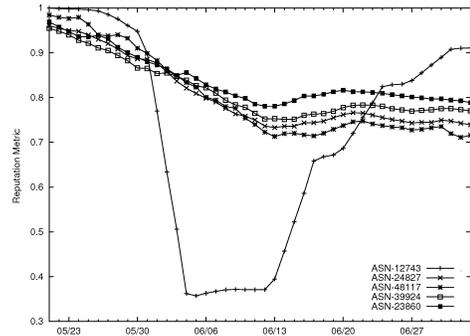


Figure 1: Example Figure/Graph