Needed Information for NSF Annual & Final Reports

Project Participants:

1. Participant Individuals
   a. Names John Hopcroft, Bart Selman
   b. Person’s involvement in the project (optional)

2. Partner Organizations - Are any organizations other than yours partners in the project, or have they been? NO
   a. If yes: Name of Partner Organization:
      Partner's contribution to the project (select one or more):
      • Financial support
      • In-kind support (organization makes software, computers, equipment, etc. available to project staff)
      • Facilities (project staff use organization's facilities for project activities)
      • Collaborative research (organization's staff work with project staff on the project)
      • Personnel exchanges (project staff and/or organization's staff use each other's facilities, work at each other's site)

   b. More detail on partner and contribution (optional, but valued by NSF):

3. Have you had Other Collaborators or Contacts? If so, describe. NO

Activities and findings:

1. What have been your major research and education activities (experiments, observations, simulations, presentations, etc.)?

2. What are your major findings from the activities identified above?

3. Training and Development - What research and teaching skills and experience has the project helped provide to those who worked on the project?

The two undergraduates, Omar Khan and Bram Kulis, who worked on the project have both gone on to PhD programs, one at the University of Texas, Austin and one at the University of California at Berkeley.

One graduate student Anirban Dasgupta, completed his PhD and is now at Yahoo Research. Andre Allavena completed his PhD and did a post doc at University of Waterloo

4. Outreach Activities - What outreach activities have you undertaken to increase public understanding of, and participation in, science and technology?

Published research in broad publication form such as PNAS and Nature.
Publications and Projects:

1. What have you published as a result of this work?

- Error bounds for correlation clustering. *ICML 2005* (with Thorsten Joachims)
- On Learning Mixtures of Heavy-Tailed Distributions. *FOCS 2005* (with Anirban Dasgupta, Jon Kleinberg and Mark Sandler)

- Journal Publications (each need:)
  i. Author(s)
  ii. Title
  iii. Journal
  iv. Volume
  v. Year
  vi. Beginning Page Number
  vii. Status of Publication
  viii. Was NSF Acknowledged?

- Books and other non-periodical, one-time publications (each need:)
  i. Author(s)
  ii. Title
  iii. Editor(s) (if publication is in a collection)
  iv. Title of Collection (if applicable)
  v. Bibliographic Information
  vi. Year
  vii. Type of Publication
  viii. Status of Publication
  ix. Is NSF support formally acknowledged in the publication?

2. Internet Dissemination

[Burt's home page]
a. If you have a relevant Web site or other Internet site, please enter or update URL(s). (Select the Web page that best represents or introduces the work and results made possible by this award.)

b. If necessary, explain or update how this site relates to the award.

c. Is NSF support acknowledged on the web site?

3. What other specific products (databases, physical collections, educational aids, software, instruments, or the like) have you developed?

a. Choose from
   - Data or database
   - Physical collection (samples, specimens, cell or germ lines, etc.)
   - Audio or video
   - Software or netware
   - Educational aid (not covered in a previous category)
   - Instrument or equipment
   - Other invention
   - Other: ____________________

b. Briefly describe the specific product that you have developed or enhanced.

c. Briefly describe how you will share this product with others.

Contributions:

1. **Contributions within Discipline** - How have your findings, techniques you developed or extended, or other products from your project contributed to the principal disciplinary field(s) of the project? **Work on agglomerative clustering made use of the instability of agglomerative clustering to define the concept of natural communities which are used to extract trends in research fields over time.**

2. **Contributions to Other Disciplines** - How have your findings, techniques you developed or extended, or other products from your project contributed to disciplines other than your own (or disciplines of colleagues and associates not covered under "Contributions within Discipline")? **The work was inherently interdisciplinary between algorithm, int sch and machine learning.**

3. **Contributions to Human Resource Development** - How have results from your project contributed to human resource development in science, engineering, and technology?

   The research effort produced two PhD’s and launched two undergraduates into PhD programs.

4. **Contributions to Resources for Research and Education** - How have results from your project contributed to physical, institutional, and information resources.
for research and education (beyond producing specific products reported elsewhere)?

5. **Contributions Beyond Science and Engineering** - How have results from your project contributed to the public welfare beyond science and engineering (e.g., by inspiring commercialized technology or informing regulatory policy)?

**Special Requirements:**

1. **Objectives And Scope** - A brief summary of the work to be performed during the next year of support if changed from the original proposal.

2. **Special Reporting Requirements** - Do special terms and conditions of your award require you to report any specific information that you have not yet reported?

3. **Unobligated Funds** - Do you anticipate that more than twenty percent of the funds under your NSF award will remain unobligated at the end of the period for which NSF currently is providing support?

4. **Animals, Human Subjects, Biohazards** - Has there been any significant change in animal care and use, use of human subjects, or biohazards, from what has previously been approved?