CrowdSearch: Exploiting Crowds for Accurate Real-time Image Search on Mobile Phones

Michael Fusaro

Multimedia Search

- Modern mobile phones are powerful
  - Most have powerful built-in cameras
- Effective search capabilities for multimedia are a necessity
- Problems
  - Image searching is a tough nut to crack
  - Video search even harder

Idea: Crowdsourcing

- Crowdsourcing: outsourcing tasks to a undefined group of people

- Improve image search
  - Humans are good at recognizing images

- How did CrowdSearch harness this?

Amazon Mechanical Turk

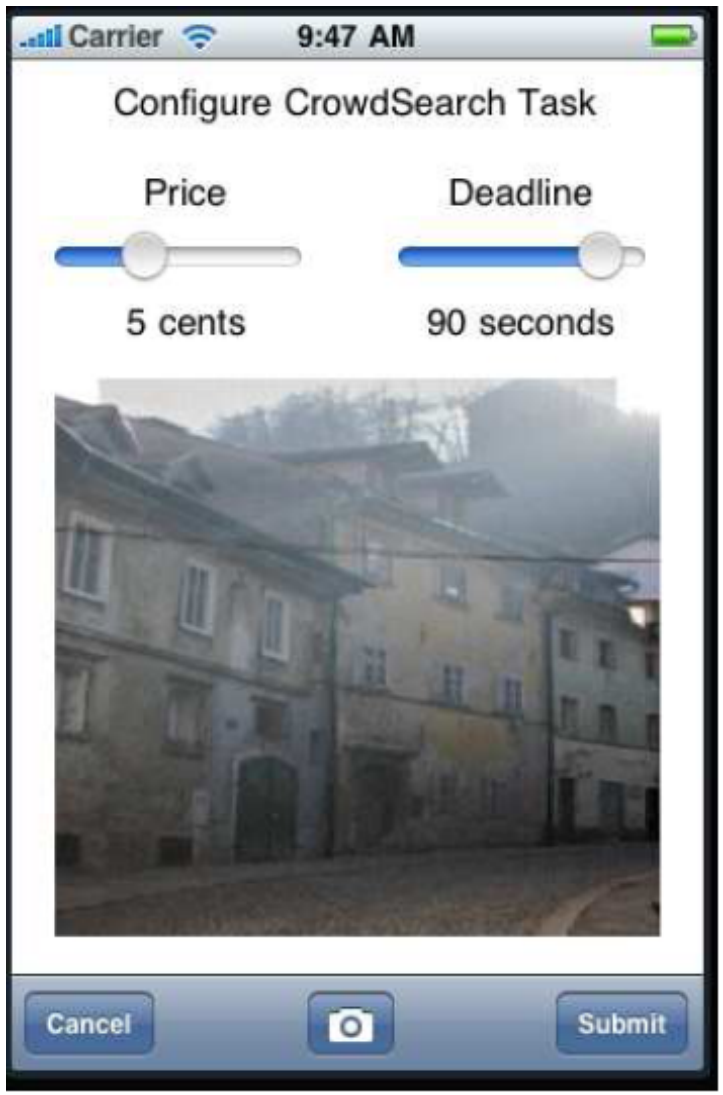
- Crowdsourcing Internet marketplace that enables programmers to coordinate tasks that are usually not feasible with a computer
- Accessible through an open API
- Users need to be paid

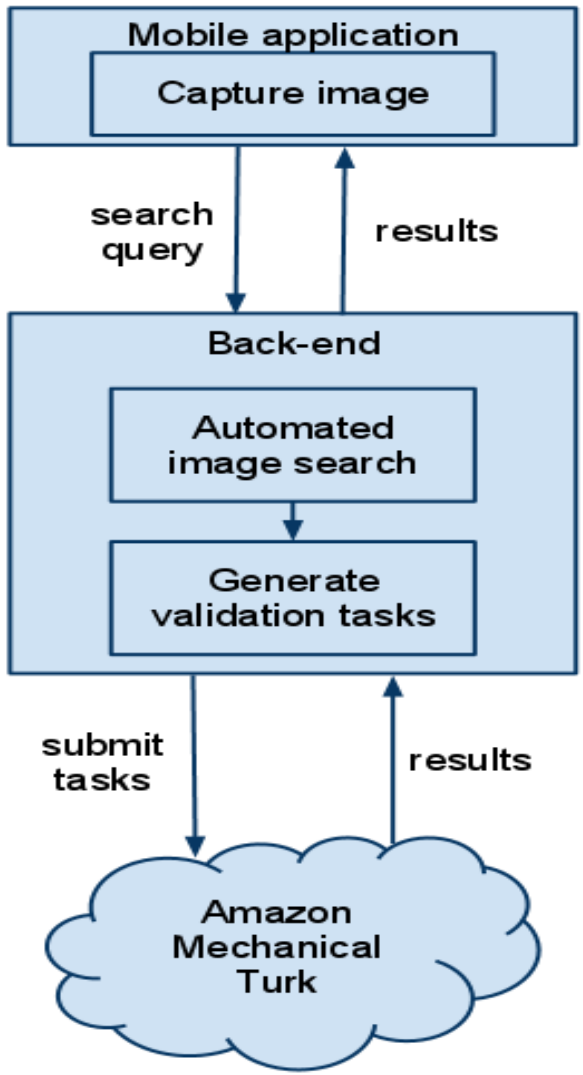
What Is CrowdSearch

- Accurate search system for mobile phones

- Consists of 3 parts
  1. Mobile phone application
     - submit queries
     - display results
  2. Back-end server
     - automated image search
     - submit AMT tasks
  3. Crowdsourcing system
     1. validate automated image search results

CrowdSearch Application

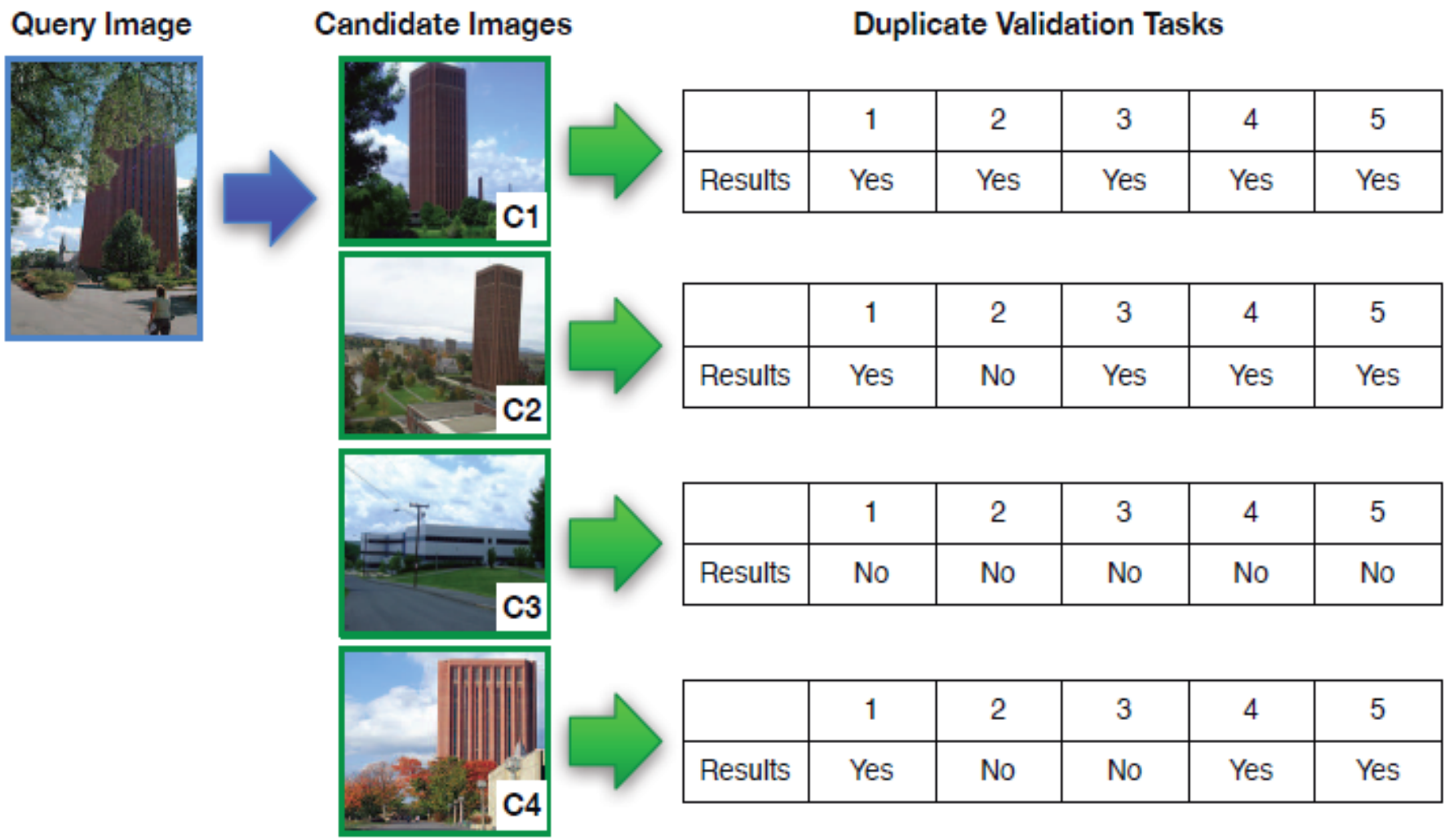




Harnessing Amazon Mechanical Turk Efficiently

- Realities
  - Tasks cost money
  - Significant delays
- Optimize for cost
  - Post tasks serially
    - pro: least expensive
    - con: takes longer
- Optimize for delay
  - Post tasks in parallel
    - pro: faster
    - con: expensive

Harnessing Amazon Mechanical Turk Effectively



	1	2	3	4	5
Results	Yes	Yes	Yes	Yes	Yes

	1	2	3	4	5
Results	Yes	No	Yes	Yes	Yes

	1	2	3	4	5
Results	No	No	No	No	No

	1	2	3	4	5
Results	Yes	No	No	Yes	Yes

CrowdSearch: Algorithm

CrowdSearch tries to strike a balance between the serial and parallel posting schemes

Goal of Algorithm

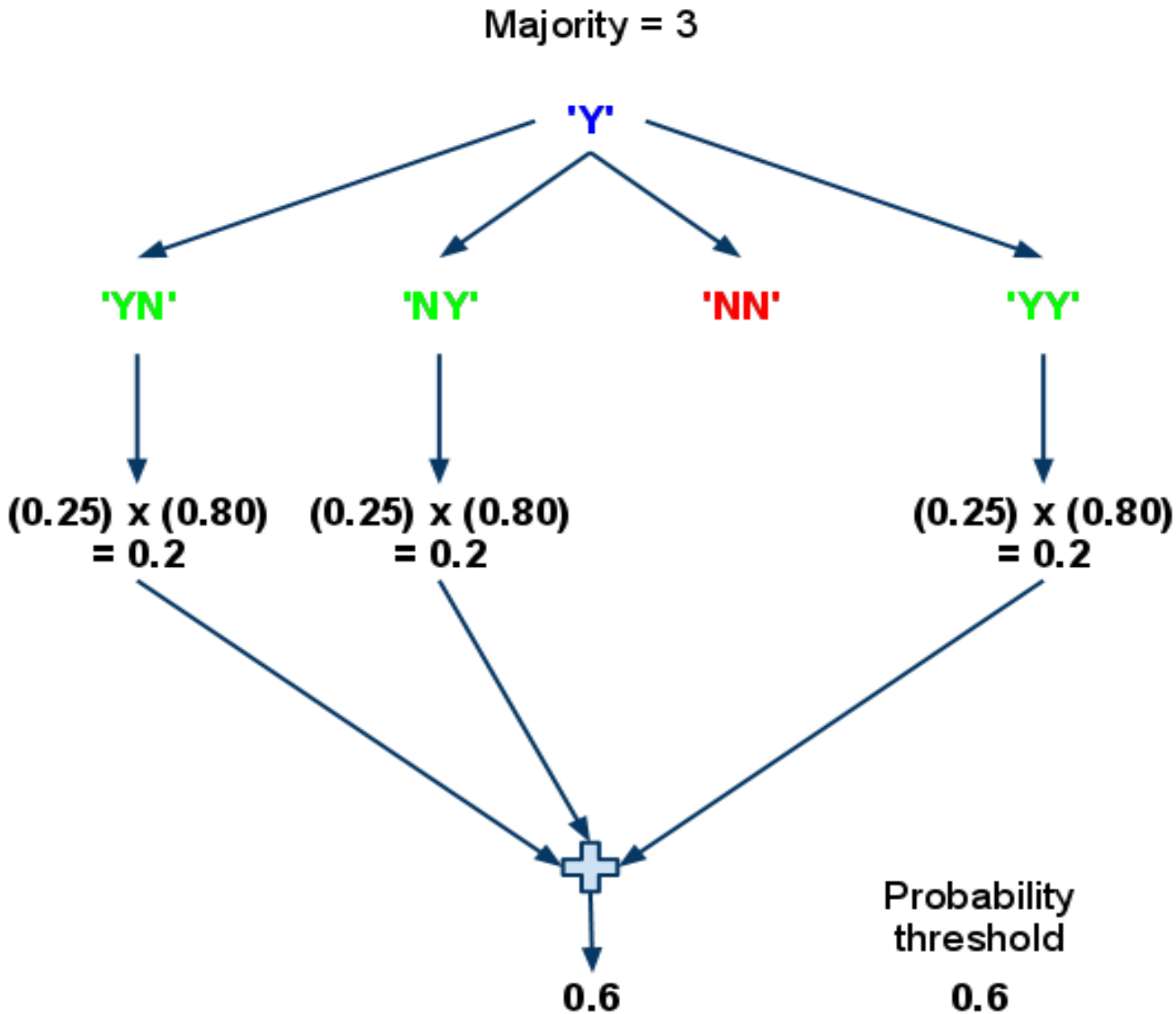
- Return at least one positive result within the predefined deadline

The Algorithm

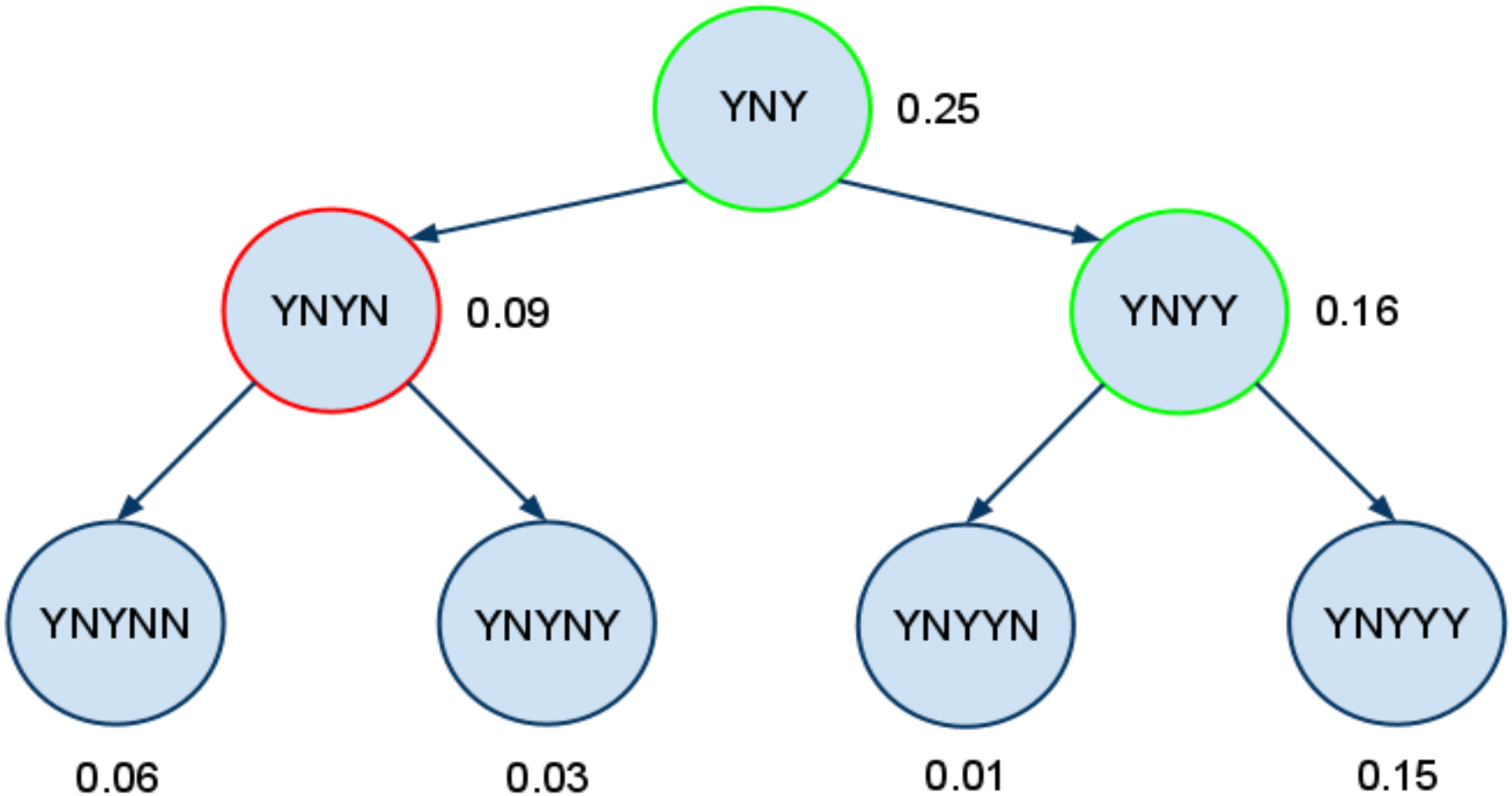
- For all current validation tasks
  - For each partial sequence received
    - Traverse all possible sequences that lead to a majority 'Yes' answer
    - Calculate probability of sequence occurring under the deadline
- If the sum of all these probabilities is greater or equal to the threshold: return true
- Otherwise: return false

- Two important functions
  - *DelayPredict()*
  - *ResultPredict()*

Example

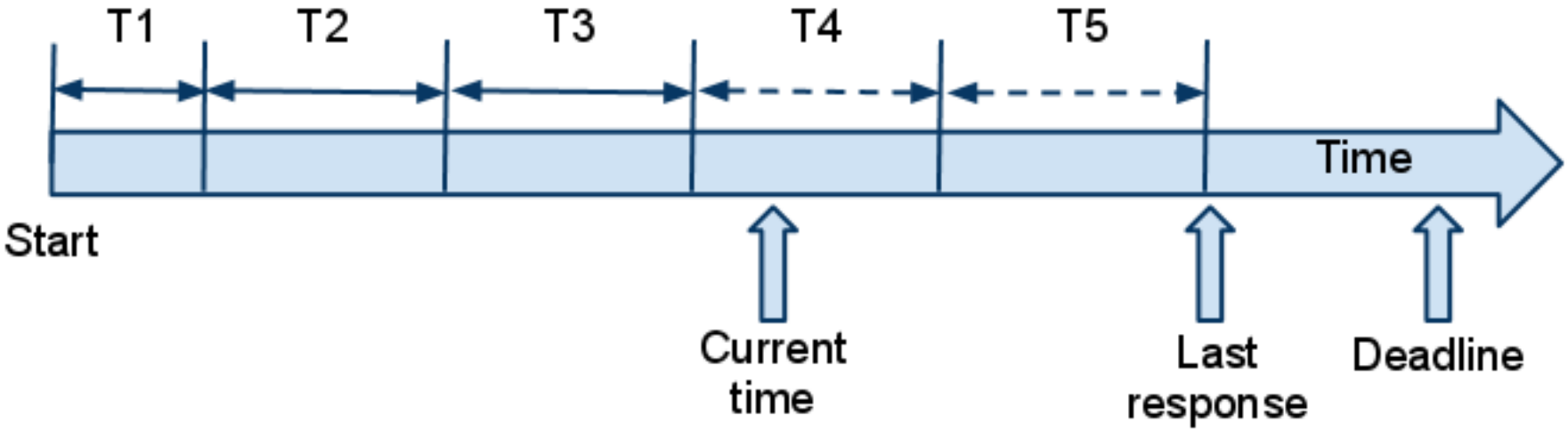

Majority = 3
'Y'
'YN'
'NY'
'NN'
'YY'
(0.25) x (0.80) = 0.2
(0.25) x (0.80) = 0.2
(0.25) x (0.80) = 0.2
0.6
Probability threshold
0.6

How ResultPredict() Works



Probability of 'YNYY' occurring after 'YNY' is 0.16 / 0.25 = 0.64

How DelayPredict() Works



- AMT validation delay has two parts
  - acceptance delay
  - submission delay

Back-end Image Search Engine

- Two major steps happen during a search

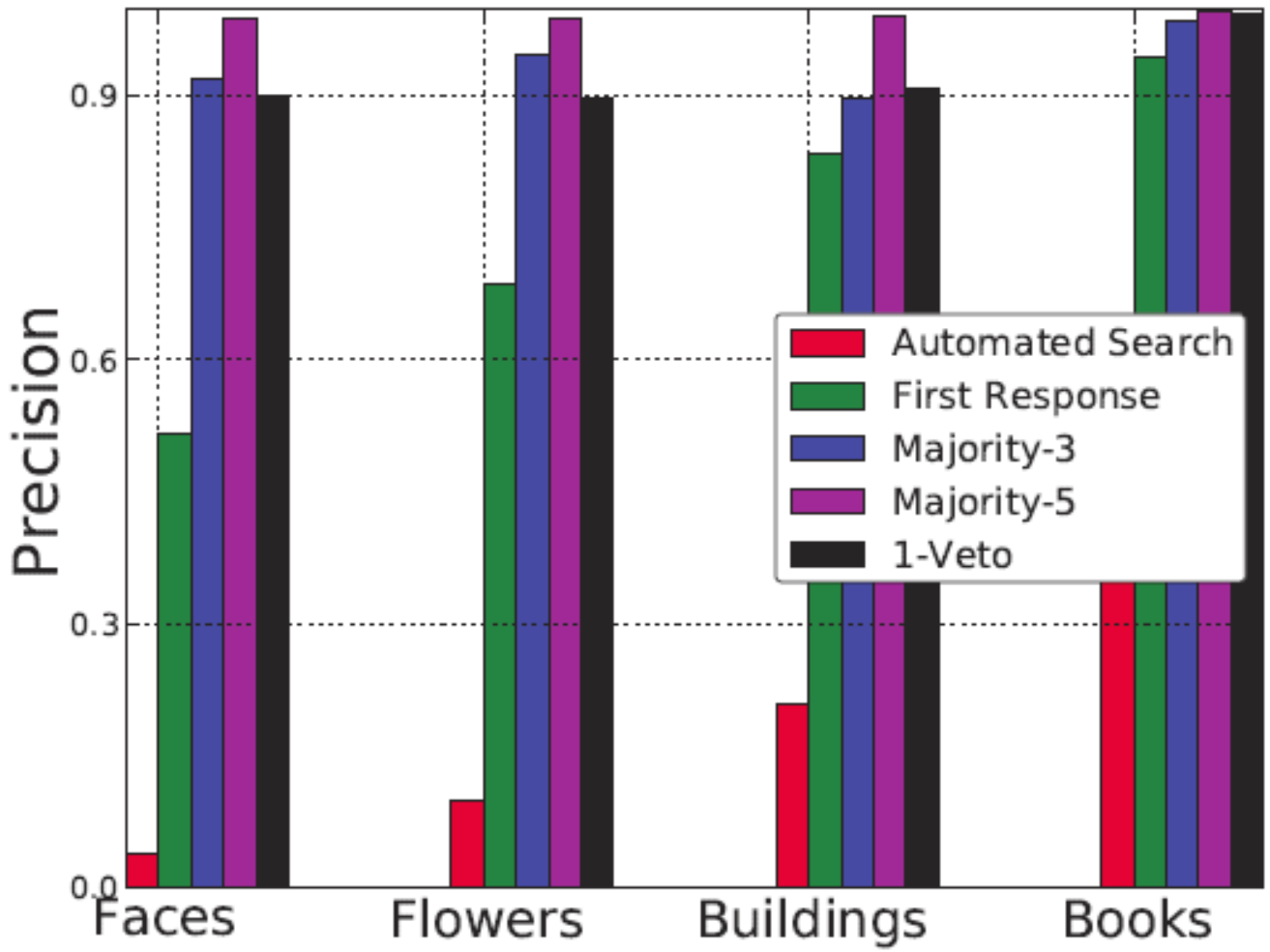
1. Extract local features from image
   - Uses a modified form of Scale-invariant feature transform (SIFT)

2. Identify closest matching image using these features

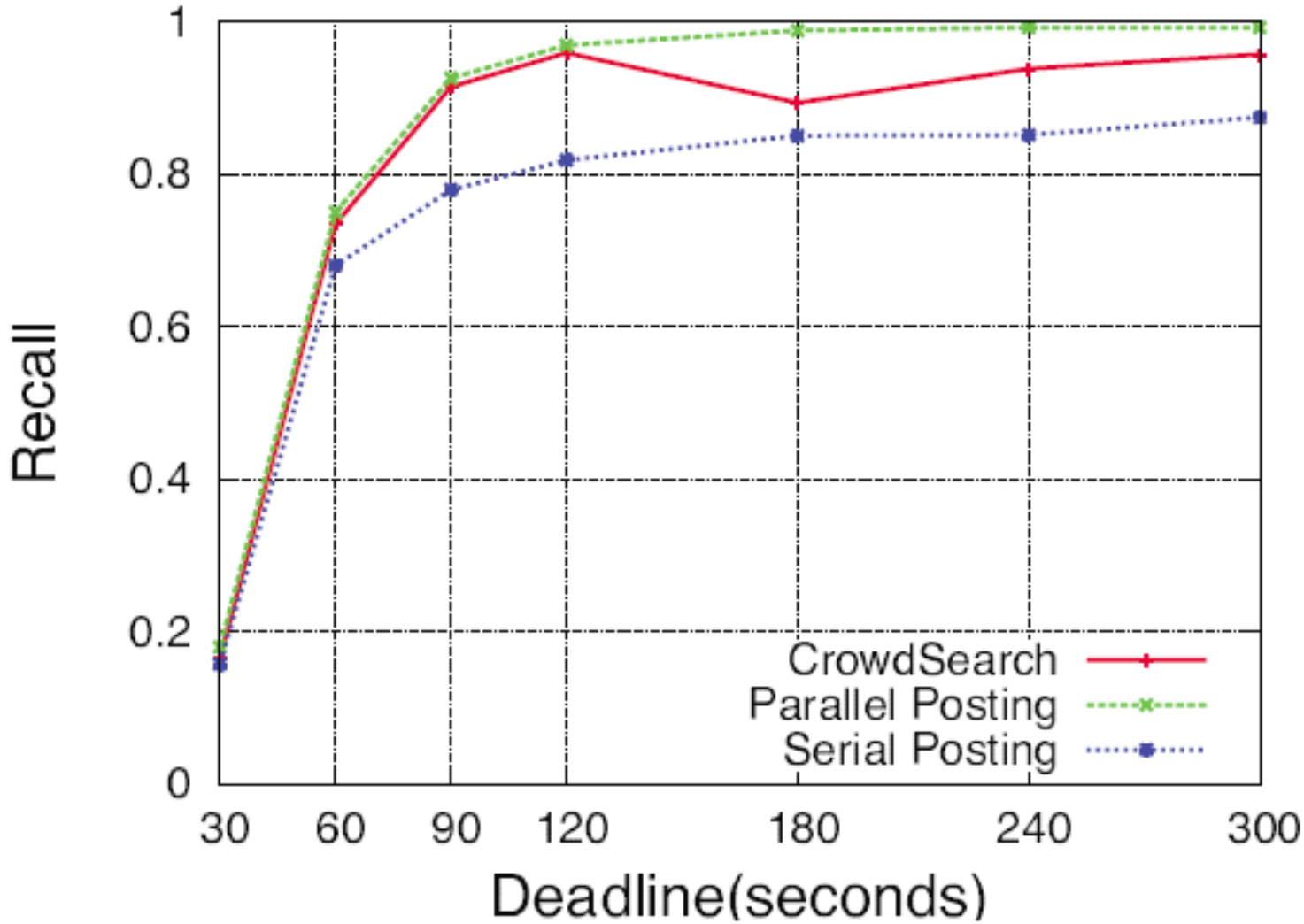
Experiment: Does it work?

- Back-end server was trained on thousands of images
- Separated into 4 categories
  - Human faces
  - Buildings
  - Flowers
  - Book covers
- 500 test images used for experiment
- Three performance characteristics measured
  - precision
  - recall
  - cost

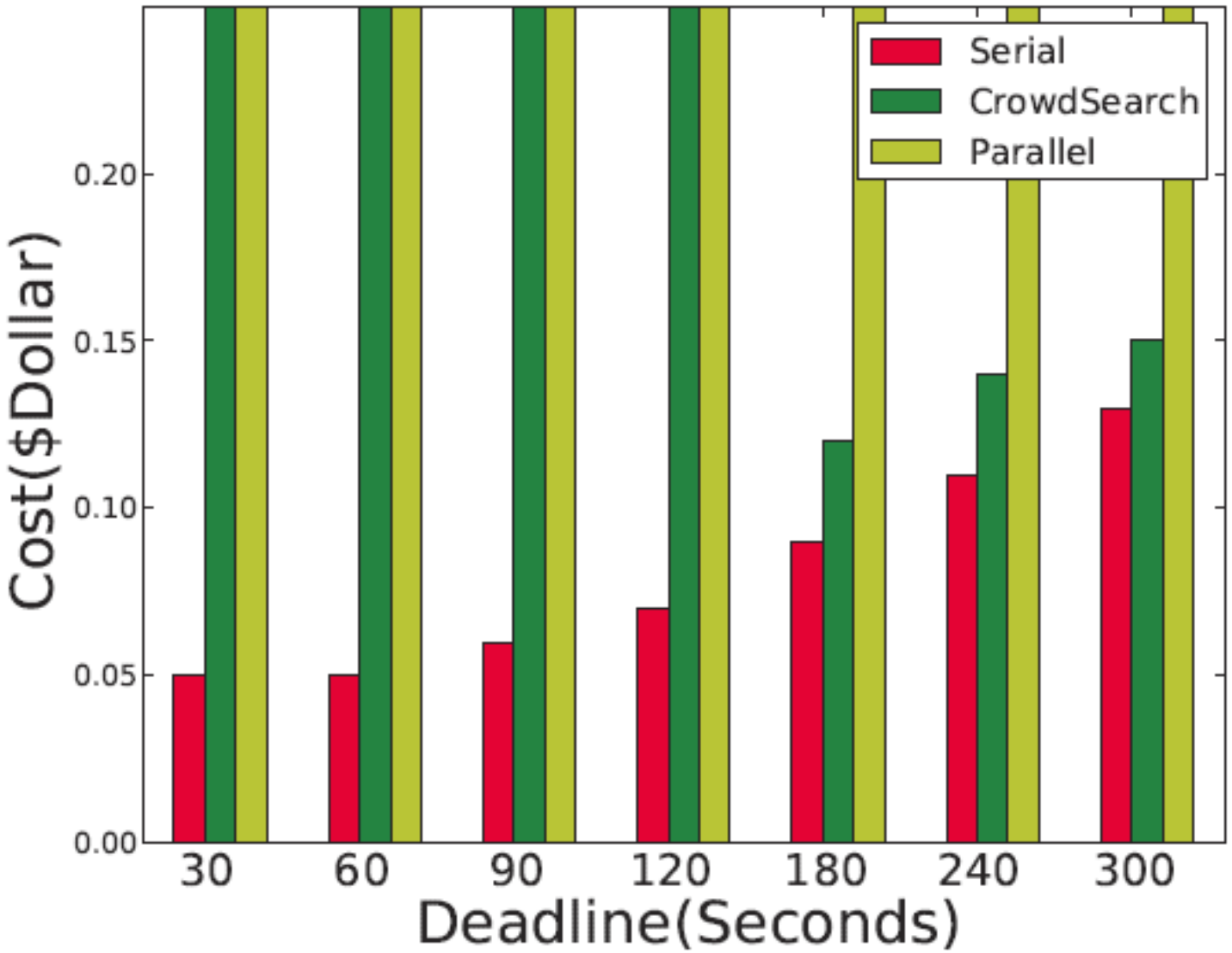
Results - Precision



Results - Recall



Results - Cost



Conclusions

- CrowdSearch algorithm was able to optimize for delay and money constraints
- Achieved > 95% search precision for several categories of images

Questions?

Bibliography

- ***CrowdSearch: Exploiting Crowds for Accurate Real-time Image Search on Mobile Phones.*** Yan, T., Kumar, V., Ganesan, D. In Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys). San Francisco, CA, June, 2010.

- ***Amazon Mechanical Turk.*** 5 February 2011. <http://en.wikipedia.org/wiki/Amazon_Mechanical_Turk>

- ***Scale-invariant Feature Transform.*** 5 February 2011. <http://en.wikipedia.org/wiki/Scale-invariant_feature_transform>