
IMC 2010 Review #34A
Updated Tuesday 1 Jun 2010 5:15:27pm EDT

Paper #34: Taming User-Generated Content in Mobile Networks via Drop Zones

==== Paper summary =====

The paper hopes to exploit the lag between creation time of user-generated content and uploading on mobile networks so that it can be scheduled at a suitable time/venue (the so-called drop zones).

Novelty: 3. Incremental improvement

==== Exciting =====

Speeding up of uploads in bandwidth constrained networks. Exploit the claimed delay between creation and uploading time to schedule the uploading at a more convenient time/place.

==== Run of the mill =====

Scheduling file transfers is not particularly interesting in general but in the mobile context it would be of value if there was any particular incentive for the user.

==== Comments for author =====

I didn't get much out of your proposal in terms of incentives for the user: why would the user want to wait till you get around to uploading it? What do they gain?

A flaw in your assumption of creation time is close to when user wanted to upload: the content may have been created earlier and given to them to upload (users uploading on behalf of a family member etc.) i.e, the decision to upload time may what really matters, not creation time necessarily. You make this assertion again in 2.2.3 ('as soon as they shoot them'). This undermines your time argument. What is the incentive for waiting once the decision has been made to upload?

2.2 It was not clear if 1.959M clients uploaded or were found in the study data?

2.2.3 Here is where you spend your time arguing about statistics of lag between creation and uploading: it would have been much more useful to

at least conduct some survey instead of simply asserting this as a fact. Especially the claim that "a bulk of user generated content" can be uploaded after a delay.

The issue of where uploads are done and the claim about battery life are another red herring: they weaken your case and do not strengthen it. If most of the uploads are done at home or work then the users are much more likely to use WiFi than 3G and besides electricity is always available (and free at work!) which renders concerns about battery life inoperative. It appears that you are stretching several non-facts to make a case for your study! At the very least you could have made a case for economic incentives: somehow users might be given additional minutes for delaying uploads or charged less or whatever. In a world of instant gratification scheduling uploads in the hope that their drive would include a drop zone so that the much loved Wireless companies could benefit, is just a non-starter idea.

That said, the rest of the paper is a standard evaluation of a delayed delivery scheme based on actual data.

Sec 4.1 Greedy v optimal: Greedy selects a larger number of Drop Zones when delivery was postponed longer so it can keep looking for zones that can deliver more content. But why does this not interfere with (or collide with) independent decisions made on behalf of other users?

Scenarios like the ones highlighted in 4.5.2 are rather speculative. We can also get more "improvement" if users reduced the amount of data they shared or if WiMax became universal.

IMC 2010 Review #34B
Updated Thursday 17 Jun 2010 11:01:38pm EDT

Paper #34: Taming User-Generated Content in Mobile Networks via Drop Zones

===== Paper summary =====

The paper finds that mobile phone uploads have increased significantly enough that they are overwhelming the 3G/2.5G network capacity significantly. Using a trace from a cellular provider, the paper finds that most users upload content from either their office or phone locations, but because users are spread around, there is a significant overload of all base stations. To alleviate this, the paper proposes Drop Zones that are locations where additional capacity is allocated. The goal is to figure out a small set of these locations to upgrade; the paper makes the observation that users can typically wait for a little time before they upload the content, thus, take advantage of mobility patterns of individual users to upload the particular mobile data at a later time when they are in the Drop Zone.

Novelty: 2. New contribution

===== Exciting =====

I think the problem of mobile uploads is an important and new problem I have not seen before.

===== Run of the mill =====

The assumption that people are going to wait to find a Drop Zone for upload is somewhat unreasonable.

===== Comments for author =====

The paper brings a fresh problem to the table, proposes a plausible solution and has decent evaluation. I like the paper overall.

I have a few comments though.

a) I am not sure of the assumption that users will wait for the Drop Zone to upload their content. Users typically are impatient when it comes to these things. Although the study points out that mobile uploads happen way after the particular picture, or video, or whatever is generated, the impatience is w.r.t .when they decide they want to do something to actually getting it done. So, it is a critical assumption that this work makes that I am not completely sure of as a business model. So, I would tone down the argument a bit.

There is a lot of repetition in section 2 that may be cut down making way for the elaboration on how it is that you decide home/work are the places where users spend most time.

b) The other thing that we are seeing that is happening is encouraging the use of Wifi networks. For example, recently, at Times Square in NY, AT&T has begun rolling out many more free wifi to alleviate iPhone-related congestion. I think that may be along the lines of what your Drop Zone might actually end up being. Have you considered this ? In which case, the range may be less, but may be much higher bandwidth.

c) One of the observations is that if users move even by 2km, it will result in significantly more coverage is not practical. I am not sure, as a user, I would drive 2 Km to drop off my mobile data.

d) Finding the optimal schedule is nice, but I would have looked at the following. Split the trace into two halves, and optimize based on the first half and evaluate on the second half. I am a little unsure that the results are going to be exactly the same. Depends on the simiarlity of usage patterns across time.

e) figure 9 could use 95% error bars in addition to the the average numbers.