EXHIBIT I

Conversation with drttol at 2005-01-24 11:45:54 on David Gucwa (aim)
(11:45:54) dr ttol: hi
(11:47:57) David Gucwa: hi
(11:48:21) dr ttol: are you working?
(11:48:34) David Gucwa: yeah
(11:48:39) dr ttol: what are youw orking on?
(11:48:49) David Gucwa: setting up the proxy part of the script
(11:48:59) dr ttol: you mean the gateway thingy
(11:49:01) David Gucwa: yeah
(11:49:13) dr ttol: to allow it to discern where to send the login info to etc
(11:49:27) David Gucwa: what do you mean
(11:49:38) dr ttol: we'll have a lot of shells
(11:49:54) dr ttol: we need the script to be able to efficiently route requests to those shells
(11:50:02) dr ttol: so that those shells execute the requests
(11:50:08) dr ttol: so it appears from different ips
(11:50:22) dr ttol: you're working on the router script essentially, right?
(11:50:51) David Gucwa: I was thinking of how I'm going to do this, and the best way I could come up with is to route all requests through the same shell, then as soon as that gets blocked, it immediately moves onto a different shell.
(11:51:02) David Gucwa: if we use all shells at the same time, they could conceivably just block them all at once
(11:51:46) dr ttol: ok, but
(11:51:54) dr ttol: we're talking about 3-10 different social networks
(11:52:03) dr ttol: 2 might block
(11:52:06) dr ttol: 8 might not
(11:53:16) David Gucwa: I'm setting up a separate list for each network. They'll all start out the same at the beginning but it will know which networks have blocked which shells.
(11:53:30) David Gucwa: Actually it would probably be better to start them off in different shells just to distribute the load
(11:53:44) David Gucwa: anyway, one particular network shouldn't be seeing more than one IP from us at a time
(11:55:30) dr ttol: andyou'll be able to catch when they block us?
(11:56:09) David Gucwa: if they firewall us out, then their server just won't respond and I catch that already
(11:56:23) dr ttol: ok, or they might feed us bad info
(11:56:25) David Gucwa: if they do something funny, like changing around the text on their page when it's our IP, then I'm going to have to make changes on the fly
(11:56:26) dr ttol: etc
(11:56:35) dr ttol: yeah, i was thinking
(11:56:40) dr ttol: we need a stable case
(11:56:46) dr ttol: like an i2hub account on each network
(11:56:50) dr ttol: that we test right now, save results
(11:56:56) dr ttol: and you have it check every 5 minutes on each network
(11:57:01) dr ttol: if the results change at any point
(11:57:06) dr ttol: it means they've changed their layout
(11:57:10) David Gucwa: that's a good idea
(11:57:11) dr ttol: or feeding us bad results
(11:57:18) dr ttol: and then the script can beep you
(11:57:54) dr ttol: depending ont he severity of the error
(11:58:01) dr ttol: whether they jus firewalled us and the fix is moving to the next IP
(11:58:07) dr ttol: or whether they changed the entire layout
(11:58:17) dr ttol: we need it working asap
(11:58:24) dr ttol: so a beeper should let you know

(11:58:44) dr ttol: we'll also need to set up alerts so like if i2hub.com goes down, or connectu.com, it'll beep you
(11:59:11) dr ttol: so each node will have a test case
(11:59:11) David Gucwa: how do you set that up, I've never heard of that
(11:59:17) dr ttol: well
(11:59:29) dr ttol: set it up so it grabs the front page of i2hub.com
(11:59:41) dr ttol: if it matches the first result
(11:59:47) dr ttol: the stable case
(11:59:51) dr ttol: then its alive
(11:59:57) dr ttol: if its not reachable
(12:00:01) dr ttol: or 404
(12:00:03) dr ttol: or whatever
(12:00:14) dr ttol: if it doesnt contain a certain key word in the htm
(12:00:15) dr ttol: html
(12:00:20) dr ttol: then you know its not the real front page
(12:00:27) dr ttol: and site is either down, hacked, unreachable, etc
(12:00:42) dr ttol: then you send an email to a special beeper email with a short message
(12:00:46) dr ttol: and it'll show up on your beeper
(12:01:01) dr ttol: "i2web down"
(12:01:04) dr ttol: "cuweb down"
(12:01:08) dr ttol: etc
(12:01:28) David Gucwa: ah, it was the beeper part that I was unsure of
(12:01:34) dr ttol: ill get one as well
(12:01:51) dr ttol: but hopefully you're good enough that by the time i call you, you would hve already left a message on my voice mail ;)
(12:02:41) dr ttol: im about 1.5 hours away right now from amherst, on a 56k meeting. i was in meetings in boston all weekend
(12:02:45) dr ttol: do me a favor
(12:02:51) David Gucwa: ok
(12:02:53) dr ttol: log into thefacebook with one of those accounts
(12:03:09) dr ttol: tell me if you're able to grab email addresses from the friends listed on that account
(12:04:08) David Gucwa: Technically possible but I think it would take a long time to do, because you'd have to descend into each of the separate profile pages for each friend.
(12:04:52) dr ttol: this is what we want for that feature:
(12:05:00) David Gucwa: also that would mean a separate connection for each friend, and they'd be logging our IP each time
(12:05:05) dr ttol: a checkbox on the importer page:
(12:05:35) dr ttol: [   ] Automatically search ConnectU for my existing friends and add them to my ConnectU account
(12:05:56) dr ttol: if checked, it will descend into the profiles and try to match the emails with existing connectu members
(12:06:24) dr ttol: if there is no match, send out an invite email, and also add the email to a master database (so we dont email the same email address twice)
(12:06:49) dr ttol: what we're creating essentially is a social network spider
(12:07:42) dr ttol: how difficult will this be
(12:07:48) David Gucwa: This one person I'm looking at has 91 friends on her list. That's going to take at least 5 minutes to crawl
(12:08:00) dr ttol: really? why
(12:09:02) David Gucwa: The problem is that it doesn't show your friend's email addresses directly on your profile page, you have to go into each friend's profile, and there's a couple seconds for each one you want to load.
(12:09:33) dr ttol: call me
(12:09:35) dr ttol: 774 230 3332
(12:09:38) dr ttol: is my number right now

(12:09:44) David Gucwa: k
(12:45:15) David Gucwa: There's a hitch in the plan to crawl through the profiles in advance
(12:45:43) David Gucwa: when logged in, I can only seem to access profiles that are on my friends list or at the same school as I am
(12:47:17) David Gucwa: so I think we're going to have to spider them on the spot when people import
(12:47:24) David Gucwa: and cache them then
(12:59:32) dr ttol: wait
(12:59:42) dr ttol: atleast we can index entire schools
(12:59:43) dr ttol: with just one email
(12:59:44) dr ttol: right
(12:59:52) dr ttol: so we just need to get you 300 logins
(12:59:55) dr ttol: or so
(13:00:03) dr ttol: to access over 1 mil profiles
(13:00:48) David Gucwa: yeah
(13:01:03) David Gucwa: if you can just give me one big file with all the logins and passwords then I can use that
(13:05:07) dr ttol: can you use the one we gave you
(13:06:12) David Gucwa: I'm not sure this is going to work actually
(13:06:32) David Gucwa: I just timed the script, and it takes about 5 seconds to fetch a profile
(13:06:44) David Gucwa: for a million profiles that's 57 days to index them all
(13:06:48) dr ttol: thats fine
(13:07:33) dr ttol: we just want to get as much as we can
(13:07:47) dr ttol: i mean, if we have all of umass
(13:07:51) dr ttol: thats great
(13:07:57) dr ttol: prob only take a few hours
(13:08:33) David Gucwa: well, except I don't just have a list of ids that belong to umass
(13:08:36) David Gucwa: I have to try each one
(13:08:52) dr ttol: what do you mean
(13:09:34) dr ttol: with one umass login and password
(13:09:38) dr ttol: you're able to see the entire school
(13:09:57) David Gucwa: If I type in the id of a profile, and that person is from umass, then I can see it
(13:10:05) David Gucwa: but there's no way to know which ids are from umass
(13:10:14) dr ttol: if you click browse
(13:10:21) dr ttol: you get a list of users
(13:10:30) dr ttol: that are from you school
(13:11:17) David Gucwa: where does it say browse
(13:11:27) dr ttol: give me a login and password
(13:11:51) David Gucwa: lbowman@wellesley.edu divya
(13:12:52) dr ttol: social net
(13:12:56) dr ttol: is browse
(13:12:58) dr ttol: it is randomly generated
(13:13:01) dr ttol: so you have to check for duplicates
(13:14:39) dr ttol: see it?
(13:15:05) David Gucwa: yeah
(13:15:09) David Gucwa: let me see how fast I can grab ids from that
(13:15:13) dr ttol: k
(13:15:18) dr ttol: and that "next" link
(13:16:06) David Gucwa: next is just the same as reloading
(13:16:38) dr ttol: k
(13:20:10) David Gucwa: I can get 3 or 4 ids a second
(13:21:07) David Gucwa: How many people from umass would you say are on thefacebook? A few thousand?

(13:21:56) David Gucwa: It'll probably take ~6 hours to index all of umass's profiles
(13:22:03) David Gucwa: and cache them
(13:24:11) David Gucwa: actually that's not true, since I can only get a random list of them
(13:24:42) David Gucwa: I can probably get a large chunk of the ids pretty quickly, but towards the end it's going to be harder to get ids I haven't seen yet
(13:25:04) David Gucwa: when does this project have to be finished?
(13:28:15) dr ttol: which project
(13:28:29) David Gucwa: the whole thing, when is it going to go live
(13:28:46) dr ttol: later this week, so we want the spider to continue to do as much as possibl
(13:29:03) David Gucwa: Yeah the spider is going to be doing the bulk of the indexing
(13:29:18) dr ttol: would two spiders indexing at the same time be twice as fast
(13:29:30) dr ttol: would three? would X?
(13:30:11) David Gucwa: if we put them all on different servers. I'm not sure how much of a bottleneck bandwidth or cpu power would be if we had X spiders running on one computer
(13:30:46) dr ttol: and one central index
(13:31:47) David Gucwa: how many proxies do we have access to at the moment
(13:31:54) dr ttol: a few
(13:31:59) dr ttol: ill order more every day
(13:32:32) David Gucwa: well I'll index what I can
(13:32:43) dr ttol: ok
(13:32:55) David Gucwa: do you have a umass login that I can use
(13:33:52) dr ttol: one sec
(13:37:58) dr ttol: rdegutis@student.umass.edu
(13:38:02) dr ttol: ruth1783
(13:41:28) dr ttol: we can have all of umass indexed in 6 hours?
(13:47:31) dr ttol: ?
(13:49:46) David Gucwa: I can have a lot of it indexed in 6 hours
(13:49:57) David Gucwa: the problem is the randomness
(13:50:13) David Gucwa: like once we have 99.9% of the id's, then we're going to almost always get repeats
(13:50:22) David Gucwa: the more we have the harder it is to get new ones
(13:50:32) dr ttol: right
(13:50:39) David Gucwa: it's diminishing returns
(13:50:44) dr ttol: so get as much as you can, then move into another school
(13:50:47) David Gucwa: k
(13:50:51) dr ttol: how many ids in 6 hours
(13:51:06) David Gucwa: At least a thousand
(13:51:29) dr ttol: tomorrow, can we set it up so that we can crawl two separate schools at the same time
(13:52:07) dr ttol: does it take 5 seconds from your house?
(13:52:25) dr ttol: would it be faster if it was on i2hub's web machine
(13:52:56) David Gucwa: probably
(13:53:12) dr ttol: can you test the difference
(13:56:02) David Gucwa: yeah
(14:03:36) David Gucwa: I'm not sure this is going to work actually
(14:03:39) dr ttol: ?
(14:03:59) David Gucwa: I'm running into problems getting the social network page to show from commandline
(14:04:59) dr ttol: well, they are reachable by browser
(14:05:06) dr ttol: so its not impossible

(14:06:32) David Gucwa: the thing about doing it from php is that it doesn't remember that you're logged in, so you can't go to the login page and then go to the search page
(14:06:45) David Gucwa: you have to go right to the search page and encode the login information into the url
(14:07:04) David Gucwa: thefacebook doesn't seem to like that for the search page for some reason
(14:08:27) dr ttol: whats a solution
(14:08:33) David Gucwa: i'm looking for one
(14:11:07) David Gucwa: you mentioned that when importing, you want to be able to send invite emails to people who aren't in connectu yet
(14:11:25) dr ttol: yes
(14:11:44) David Gucwa: I'm not sure how vital that is, but if we don't add that functionality then we will not need to crawl into other people's profiles and importing will take 5 seconds instead of 5 minutes
(14:12:18) David Gucwa: we can still get a list of their friends by id and find those friends on connectu if they have profiles already
(14:12:23) dr ttol: well
(14:12:25) dr ttol: if we crawl now
(14:12:31) dr ttol: then we'll also know who they're friends with
(14:13:01) David Gucwa: I think it would be a lot less trouble to do it at import time
(14:13:06) dr ttol: so if we have software that can crawl through every profile on a school, we'll know all the social networks
(14:14:04) dr ttol: there are two approaches to this
(14:14:09) dr ttol: we can either:
(14:14:20) dr ttol: a) index through "social net"
(14:14:37) dr ttol: b) index at import time, and import allt heir friends and all the profile information of themselves and their friends
(14:14:44) dr ttol: both will likely have the same outcome -- an index of the entire net
(14:14:51) David Gucwa: I'm not sure why we need the profile information for their friends
(14:14:53) David Gucwa: at import time
(14:15:54) dr ttol: because when that friend signs up, we already have their profile information
(14:16:53) David Gucwa: we can go fetch it pretty quickly when that friend signs up
(14:18:05) dr ttol: we still need the email addresses
(14:19:02) David Gucwa: That's what I was asking, how important is that functionality
(14:19:10) dr ttol: pretty important
(14:19:38) David Gucwa: which is more important, a short import time or sending email invites to non-users
(14:19:52) dr ttol: cant we send the email invites later
(14:20:06) dr ttol: it'll still be short import time, we can just have another software do the crawling of non-users
(14:20:18) dr ttol: we want to be able to send friend requests for email addresses already in our database
(14:20:20) David Gucwa: is the invite coming from connectu or from the person who just signed up
(14:20:30) dr ttol: connectu
(14:20:36) David Gucwa: okay that's fine then
(14:20:46) dr ttol: but via importer software
(14:20:53) dr ttol: like a sendmail()
(14:21:12) David Gucwa: right
(14:21:25) dr ttol: but we still need to crawl

(14:21:28) dr ttol: to the profiles
(14:22:05) dr ttol: the emails should be saved in a database
(14:22:14) David Gucwa: but the user doesn't need to wait for that, right? there's nothing in the crawled friends' profiles that the user needs for his own?
(14:22:14) dr ttol: so before the sendmail(), it checks to make sure the email isnt in the database
(14:22:31) dr ttol: because there will be common friends, and we only want one email
(14:22:40) dr ttol: right
(14:22:46) David Gucwa: ok
(14:22:50) dr ttol: well
(14:22:54) dr ttol: except the friend requests
(14:23:17) David Gucwa: okay but that means the friend already has a profile with connectu
(14:23:20) dr ttol: that part of the software needs to crawl through their friends list and check to see if they're already on connectu
(14:23:30) dr ttol: if they are, send a friend request through connectu software
(14:23:31) David Gucwa: so we don't need to crawl their facebook account
(14:23:46) dr ttol: if they're not, add email to database of emails sent, and then also do a sendmail()
(14:23:58) dr ttol: if email is already in database, bypass sendmail
(14:24:02) David Gucwa: as long as we store their facebook id on their connectu account, we can look them up on connectu without crawling their profile on thefacebook
(14:24:23) dr ttol: ok
(14:25:05) dr ttol: the point is, if im a user, i will import my facebook account, and i expect the software to look for my friends automatically and send requests out if they're on connectu, and send invite emails out if they're not
(14:25:38) David Gucwa: does the user get an option to send out invites or not
(14:26:10) dr ttol: no
(14:26:13) David Gucwa: good
(14:26:14) dr ttol: its part of the first checkbox
(14:26:28) dr ttol: if they want auto friends, auto invites is part of the functionality
(14:26:35) David Gucwa: that makes things a lot simpler
(14:26:57) dr ttol: but you still need to gather the email addresses from the profiles
(14:27:43) David Gucwa: right
(14:28:55) dr ttol: im just thikning that if we crawl through the entire facebook school by school now
(14:29:09) dr ttol: then we essentially dont need to hit their site during impotr time
(14:29:15) dr ttol: because we'll have all the info locally
(14:29:25) dr ttol: we can just verify their login info with facebook
(14:29:36) dr ttol: and if succesful, the user will claim ownership to a certain profile we stored locally
(14:29:47) dr ttol: and all the friend requests will be sent out
(14:29:52) dr ttol: (whether through connectu or via email)
(14:29:54) David Gucwa: I don't think it's feasible to crawl their entire site before ours goes live
(14:30:38) dr ttol: ok, but if you had a friend, Adam Wood
(14:30:41) dr ttol: and I had the same friend
(14:30:47) dr ttol: that means the importer will hit thefacebook twice
(14:30:50) dr ttol: same profile
(14:30:59) dr ttol: and if theres a third friend that also had Adam Wood listed
(14:31:00) David Gucwa: the importer will cache any profile it crawls through

(14:31:01) dr ttol: it hits it again
(14:31:09) dr ttol: okay
(14:31:13) dr ttol: so only one hit per profile
(14:31:15) dr ttol: right?
(14:31:16) David Gucwa: yes
(14:31:19) dr ttol: ok good
(14:31:25) David Gucwa: I just don't think we can expect to do them all in advance
(14:31:31) dr ttol: ok
(14:33:21) dr ttol: but it may be good to have, at some point, software that caches, or do you expect the importer to do all the caching we need
(14:34:28) David Gucwa: I was expecting to write the caching for just the importer
(14:34:48) David Gucwa: the info will be there for any other reasons we need it
(14:35:08) dr ttol: ok
(14:35:10) dr ttol: eta
(14:38:03) David Gucwa: for spidering/caching on import, the rest of the day probably
(14:40:03) dr ttol: ok, what about friends
(14:40:21) dr ttol: checking to see if friends are there, and triggering a friend request event
(14:40:36) dr ttol: and db of emails that we already sent invites to
(14:40:59) David Gucwa: do you know how easy or hard it is to add a friend to a connectu profile
(14:41:05) dr ttol: ask joel
(14:41:12) David Gucwa: It would be pretty easy to just pass some information to an existing script if one exists
(14:41:22) dr ttol: im not sure which is better
(14:41:25) dr ttol: to restrict one invite per email
(14:41:27) dr ttol: or no restriction
(14:41:41) dr ttol: "One of our members, David Gucwa, has just invited you to join him in ConnectU"
(14:41:50) dr ttol: "One of our members, Wayne Chang, has just invited you to join him in ConnectU"
(14:41:59) dr ttol: someone may get 20+ emails
(14:42:06) dr ttol: that might not be bad
(14:42:08) David Gucwa: that could get pretty obnoxious
(14:42:20) dr ttol: ok
(14:42:22) dr ttol: so one email per user?
(14:42:30) David Gucwa: well I don't know if obnoxious is what you want
(14:42:36) dr ttol: no
(14:42:39) David Gucwa: it would certainly be a good incentive to join connectu and make the invites stop
(14:42:39) dr ttol: we dont want to be evil
(14:42:41) David Gucwa: heh, okay
(14:42:58) David Gucwa: I can remember each email then
(14:43:03) dr ttol: ok
(14:43:26) dr ttol: where is your home town?
(14:43:38) David Gucwa: sandwich, on the cape
(14:44:17) dr ttol: ok
(14:44:41) dr ttol: there is going to be a cambridge office opening soon. i want amherst to be the better one
(14:45:46) David Gucwa: ok
(14:46:07) dr ttol: so everything will be finished tonight?
(14:46:45) David Gucwa: everything for the spidering part
(14:46:51) David Gucwa: I still have a bunch of other work to do
(14:46:59) David Gucwa: I have to make all the other importers

(14:47:13) David Gucwa: speaking of which I need logins for the other sites
we're importing from
(14:48:50) dr ttol: right
(14:49:06) dr ttol: ill get that to you, but we need to finish thefacebook
importer first
(14:49:12) David Gucwa: ok