LINKEDIN HASHDUMP AND PASSWORDS

Unless you have been living under a rock (not judging, just that you may not get wireless there) you should have heard about the 2012 LinkedIn data leak. The hacker has released about 6.5 million hashes. So, to start things off, the hashdump is here. Ok, go ahead and unzip that. What you will see is that the dump has 6,458,020 entries. Now somethings you will notice when inspecting this file:

- The hashes are SHA1, unsalted
- There are no repeated hashes. Probablt means that this list was already sorted and optimized for cracking by the hacker
- There are a bunch of hashes starting with 00000. Current thinking is that these are the ones already cracked by the hacker and they are trying to mask those hashes.

So first things... WTF LinkedIn!!! Agreed and moving on...

Once you have unzipped the zip file, you will want to sort the list into masked and unmasked hashes. Being a linux CLI snob I would suggest something like:

```
> cat ./combo_not.txt | grep -E "^00000" > masked.lst
> cat ./combo_not.txt | grep -vE "^00000" > unmasked.lst
```

Now the reason for doing this is you are going to download hashcat (homepage here), but the special version is the one you want (see here). The reason you want this special version is because it has a tweaked MD5 variant specifically aimed at those masked hashes. You see, even with the first 5 characters masked the SHA1 hash is pretty much unique, and with the mask being all "0", it makes it even easier to tell which are masked. So if you take a look at those 2 seperated lists youwill see there are 3,521,180 masked entries and 2,936,840 unmasked entries. When you use hashcat, use "-m 150" for the masked hashes and "-m 100" for the unmasked hashes.

Now none of this is that new, but I think this list is very important since this is an actual verified list of user passwords. Not a dictionary of "could be" but a list of "has been". This means from a reuse point of view it is very useful. For that reason I am putting up the list of hashes I have already cracked. I do not have a dedicated cracking rig or GPU's or Amazon or such, this is just little old me plugging away at the list. Get it here. (link now updated as below)

That list is a snapshot of where I am in the cracking process, specifically:

```
masked (150) = 2746578 of 3521180 / 78% done (774602 left)
unmasked (100) = 852784 of 2936840 / 29% done (2084056 left)
total = 3599362 of 6458020 / 55% done
```

So if you have not started cracking the linkedin hashes, using that list will get you to where I am at least. I am still working away at it and as I get significant updates I will post those updated lists.Grab the files and have fun.

Update - 9 September 2012

Ok, I have made some large gains:

```
masked (150) = 3108522 of 3521180 / 88% done (412658 left)
unmasked (100) = 1647836 of 2936840 / 56% done (1289004 left)
total = 4756358 of 6458020 / 73% done
```

The updated dictionary is here, and the specific non-standard rules that have worked for me is here.