ॐ श्री गोरक्ष

PSOSM

(Notes)

Broad topics

- Overview
- Linux / Python / Twitter API / MongoDB / MySQL
- Trust and Credibility
- Privacy
- Social Network Analysis. NLTK.
- e-crime
- Plotly / highcharts / Geolocation analysis.
- Policing
- Identity resolution.

week 1

1.1 nothing to note

1.2 social network of diff. categories (diff data generated).

- Youtube (video sharing)
- Flicker, imgur (image sharing)
- foursquare (location based)
- linked In (professional connection)
- instagram (has follow request based approch) unidirectional

Type of social network.

1) Ephemeral - Content is posted and it destroys itself after some time.
Ex - Snapchat / Whisper isn't ephemeral more Tinder or Pintrest

2) Anonymous S.N. - (identity is not known)
Ex- Whisper

4/5 V's of online social media

1) Velocity: Speed at which data is generated on SM.

in 60 sec, 400 hrs of video uploaded on YT.

2) Variety - Variaty of content (images, blogpost video etc)

3. Veracity - to find out information posted on SN is truthful or not (Vry hard).

4. Volume - size of content generated

5. Value - we can have above 4 V's but if data isn't of any value then its useless.

Another type of classification of S.N

→ bidirectional → (to be friends both have to accept) (data flows in both dirn). (linked In)

→ unidirectional → (twitter) anyone can follow anyone (data flows from one to another & may be otherway around).

twitter (trends) what's happening around.

perception: whatever we see on S.M. is not always true

week 1.3.

- S.M. in disaster management crisis
- first time in 2009 - twitter - plane landed in Hudson river.
- student committed suicide (cyberbullying)
- find missing child (viral posts)
- great help in Nepal Earthquake.
- Egypt revolution
- Fake news on Boston Bombing
- fake news through hacking verified news accounts.
- Hurricane Sandy: fake images

Tutorial Ubuntu:

apt-get - c.l. package management tool.

Response code	meaning
200	OK
403	Forbidden
201	Created
204	no content
401	unauthorized
500	internal server error

Week 2.1

Data collection through API (Application Programming Interface)

facebook Graph API returns in JSON (JS object notation)

phpMyadmin - looks data through MySQL (helps)
Robo Mongo - helps look data through MongoDB.

week 2.2. (Trust and Credibility on O.S.M)

General flow of things in times of disaster on OSM

- fake news being generated very fast / rumour
- true ~~new~~ news comes very much after the fake news being spread

Tutorial 3 FB Graph API

curr version 2.10.
requests permission from user for each info access

Week 3

3.1 Misinformation on Social Media

misinfo on SM results in:
- chaos
- effect on share market
- rumour

spikes in data collected on SM. is corellated with spikes in real world incidents

fake accounts in themselves have closed community & connected closely

TweetCred Trends API / Streaming API ↔ DB → feature set generation → result.

FB inspector - a chrome extension for fake news detection.

FB Graph API → ground truth extraction → Gene. feat. vectors → Supervised learning → Rest API

3.2 Privacy: has different meaning for different persons in different scenarios.

Westin's 3 type of category of people

- Pragmatists (60%): depending upon situation respond. share details for benefit
- Fundamentalists (15%): who have very strong privacy expectations.
- Unconcerned (15%): who will give away imp infor. (personally identifiable) info. easily for minimal returns

Although internet makes you anonymous but it is easily possible to identify who you are on internet by looking at your browsing behaviour

Privacy India 12 survey

- Most. People are going to accept friend request of opp. gender

- Although ppl are concern about privacy but feel confident after applying privacy setting to their profile.

Tutorial 3. Twitter API / MySQL, MongoDB. taking a rest here.

Week 4.

Privacy and Pictures in OSM.

Privacy is the claim of individuals, groups or institutions to determine for themselves, when how and to what extent, information about them is communicated to others

. each individual is continually engaged in a personal adjustment process in which he balances the desire for privacy with the desire for disclosure and communication

Forms of Privacy.

. Information → Internet
. Communication → Telephone
. Territorial → Living space
. Bodily → self (CCTV camera)

- in 2000, 100 Billion photos were shot worldwide
- in 2010 2.5 Billion photos uploaded by FB users inmonth
- in 2015, 1.8 Billion photos uploaded on FB, WP, SP etc everyday

- FB/MS/GGL acquired licensed products that do face recog

colluding things

- Increasing public self disclosure through OSN.
- Improving accuracy of face recog
- Re-identification techniques are getting better

Latanya Sweeney combined Medical data with Voters' list.

- online - online - pittpatt + match.com + FB data → 10% accura
- online - offline - stalking in campus - 38% accura
- 4 digits of SSN 27%

Week 5 (Policing)

5.1 from publically available data it was possible to reidentify people on ~~match.dot~~ match.com and facebook

How police can use OSN.

#my MPD backfired

prob: fake police org accounts (identified by low acitivities or not verified)

Can OSN support police to get actionable info about crime & residents opinion about policing activities

methodology.

- content
  - missing human
  - Query
  - Traffic
- style
  - formal
  - informal
- Type
  - Acknowledge : say Thanks
  - Reply : suggest solution
  - Follow up : Ask further questi
  - Ignored : no reply

5.2.

Actionable Info ? = no police at XYZ road, area not safe

5.3

measuring human behaviour.

citizen to citizen , C to pol, pol to C , P to P

P_{PLC} - police posting & both C & P interact

measure of behaviour: Topics ngram / Kmeans clustering

Emotional LIWC or Anew dictionary

valance : way of saying +/- ve things

Arousal : intensity of posts.

if police takes part in discussion : anxiety is low & discussion closes faster

Aim :

1) Help communities to make consensus based decision regarding support and action they seek from police

2) Help gauge changing emotions & behaviour among citizens.

Week 6

6.1 e crime on OSM.

- Phishing
- Fake X
- Fake follower / likes
- clickbait
- Account compromise
- Acc. impersonation
- work from home scam

Phishing: Act of tricking someone into handing over her login credentials in order to exploit personal information

(plz click to accept new T&C)
(update pan)

Whaling: targeting big companies CEO.

fake reply about help to frustated customer.

fake live stream / discount / online survey / tip / reviews

social reputation → no of likes / friends / follows → more influence in society

← social status →

also

clickbaiting: getting you to click on links which are not legitimate.

#hijaking: using hashtag for purpose other than intentionally used.

compromised accn - hacking acc of ppl with # ?? follows

impersination - imposter.

week 6.2

Pagerank. idea where websites exchange reciprocal lengths with other websites which are not there otherwise to improve rank.

link farm is a form of spamming the index of search engine (spamdexing, spamming)

why link farming in twitter:

- Twitter = web within web.
  vast amount of info in real time
  - Search engine rank users on follower rank,
  - Pagerank to decide whose tweet to return as search result.
  - High indegree (# followers) seen as a metric of influence
  - Klout score influenced by twitter indegree
    ↳ rate users on O. Social influence.

link farming Twitter: spammer follow others users and attempts to get them to follow back (reciprocity)

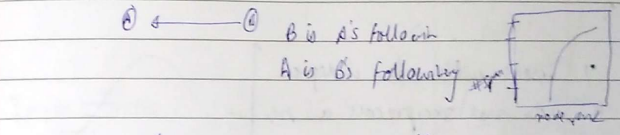
"following back is considered social etiquette".

8% of total 25 mi links posted on twitter are spam

click rates is high 0.13%.

spammer: 82% of spam followers overlap with spam ~~tgt~~ target.

16% of active acc display high value of automation most automated account spoof as browser (web)



within 10^4, 7 spammers if listed acc to node rank
~~$2 \cdot 10^4$~~ 2×10^4, 2K spammers

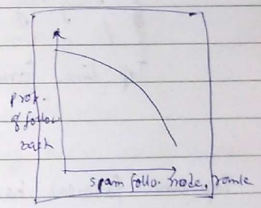
so spammers are popular than thought.

Twitter ~~x~~

Week 7.1

Spammers.

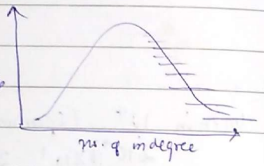
1) Top 10^5 spam followers account for 60% of all links acquired by spammer



follow back?

Probability of response vs indegree

- users with low indegree do not reciprocate a pop link to spammers



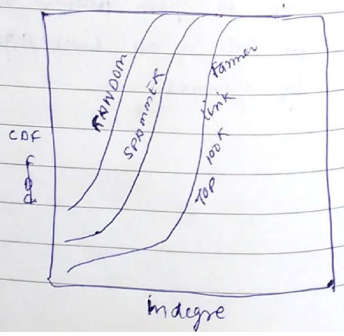
- Responsiveness increses with increasing no. of indegree

Top 5 link farmers : acc to

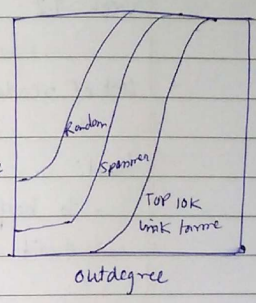
# links to spammers	Pagerank (They will follow you))
Lenny weiss	Barak Obama (election)
Judy Rey	Britny spear

Top link farmers are not spammers.

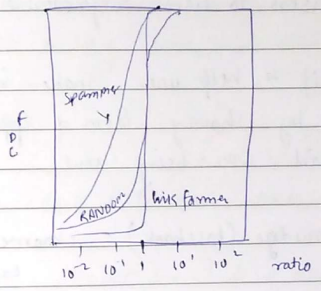
Top links farmers have very high indegree compared to spammer & random sample



Top link farmers have high outdegree compared to random or spammers



In/out ratio.



The value to top link farmer is nearly 1

conclusion : character ss of link farmers

1) legitimate, popular, highly active users engage in link farming.
2) increase social capital & influence

x-7.1-x

link farmers (indegree) >> spammers, random

" (outdegree) >> spammers, random

week 7.2.

- Cost of reading privacy policies ~ 781 billion $. per year

Goal:
- to help user make informative decisions.
- avoid users to have regrettable online disclosure.

Facemail (MIT) → help user make informative decision by showing faces of ppl to whom the email was being sent.

Picture ~~nudge~~ nudge (facebook) - chrome browser extension
showed pictures of people who will get to see this post & other information

Timer nudge: ten seconds delay in finaly post.

Sentimental nudge: Show how other ppl may perceive your post.

Analysis metrics.
1) no. of change in online privacy settings
2) no. of cancelled or edited posts
3) posts frequency
4) Topic sensitivity

↳ intervention help users make betts decisions
↳ still a lot of work to be done

-x-

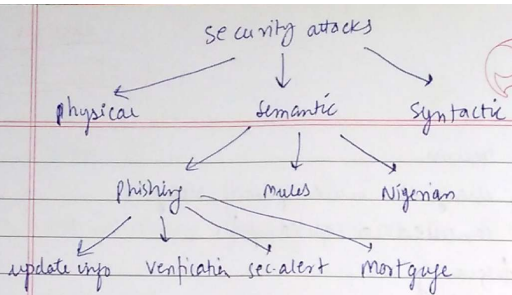
week 7.3.

Phishing attacks / Semantic attack.

semantic attacks are kind of attack where humans are targeted.

3 type of attacks → physical / syntactic / semantic.
phy access to PC ↑ denial of S., buffer overflow ↑ attack the way we assign meaning to content (phishing)

semantic barrier b/w machine & human
machine thinks other than human
(hard to fix this problem)



Types of phishing attacks -

- Phishing
- Context aware phishing / Spear phishing → email to students take course (they know something)
- Whaling - sent to CEO
- Vishing - over phone
- Smsishing - over SMS.
- Social Phishing. ???
  using publically available database to exploit you
- Indiana university: investigation

1. Public data / Blogging Social networks
2. Social Network database
3. email sent bob by eve ~~Alice~~ (spoofed as Alice)
   (check this link)
4. Email opens, redirected to another website.
   enter username & password.
5. web & authentication logs.
6. goes to authenticator.

Success	Failure
server overload try again	Authorization failed try again

Gender analysis: phishing more successful if it came from opp gender.

Reaction Anger / Denial / Misunderstanding / Underestimation.

Conclusion - extensive educational campaign
- Browser solution
- Digitally signed email.
- OSM provides info for successful attack

week 8.1 (Profile linking).

- Difficulty in profile linking.

problem of de-duplicating audience

Challenge : Heterogeneous OSM. diff type of info in diff. OSM.

Attribute evolution - values changes over time

- approach - list common attribute
  - compare attribute value using syntactic, semantic or graph based method.
  - High similarity denote same profile
  - check most current attribute.

why only username -

- unique attribute to user
- universally & publically available attribute
- Homogeneous, character & length restricted
- Easier history collection

methodology: Twi → feature set → labeled dataset → supervised classifier

FB →

Ground truth collection -

1) self identification behaviour of users.

supervised classification

1) independent supervised framework
2) cascaded " "

highest accuracy - (Exact best match → SVM (linear)

Conclusion:

1) Profile linking may be necessary for many organisation
2) Better profile linking is possible with past history of user handle.

week - 8.1. -x-

Week 8.2.

1) Anonymous Network - 4chan
Whisper
Secret
YikYak
Wickr

Why use Anonymous network?
- increasing awareness of privacy
- Snowden disclosure
- PRISM surveillance prog.
-

Terminology / Claims whispers
Replies
Anon names
Does not associate pid, history
Heart a message anon
Private msg

Data collected through Whisper network

Whisper ID, Timestamp, plain text of whisper, Aut. nickname & location tag, # of replies, likes.

1) 55% of whispers receive no reply.
2) 25% have chain of at least 2 replies

3) 54% of replies arrive within 1 hr. of post
4) 94% arrive in 1 day

5) If a whisper does not get attention shortly after posting it is unlikely to get attention afterwards.

Post per user

80% user post less than 10 total whispers or replies.
15% post only replies no ~~post~~ whisper
30% post whisper, no replies

network analysis
avg degree = 9 in comp to FB (190) & Twi (39)
↑
interact with more no. of ppl

clustering coeff low - means don't club ppl & which comp comm. to each other
(Random graph)

Content Moderation:

- 18% of content deleted in comp to Twi (4%)

Keyword most related to deleted whisper
sext, wood, mayphy., selfie, send, f, dm, pm, chat.

Keyword least related to deleted whisper
Emotions, Religion, Entertain. Work., lifestory.

70% of deleted whispers are deleted within 1 week after posting

User interaction:

90% of 2 users are located with in same state
75% have distance < 40 miles

although daily new users are adding up but still network traffic is stable, ppl disengage

week 8.2 ~~X~~

Week 9.1.

Privacy in location based Social Network
9GAG, LLR, FB, YELP, 4square, Twitter

Foursquare
- check in (only shared with friends)
- mayorship, tips and dones with everyone

- one can easily infer homecity of around 78% of ppl with 50 km error
- mayorship - person who is being to that place most no. of time in last 60 days.

most of data is in research ppr. (to read from over there)

Cronbach alpha: measure of internal consistency, how closely related a set of items as a group
> 0.8 ~~Excel~~ Good, > 0.7 Acceptable, > 0.6 questionable, 0.5 poor

NDCG: normalised discounted cumulative gain (quality of ranking)

MongoDB is NoSQL DBMS. | collecting ground truth → crowdflower
→ Amazon mechanic turk

we know where you live : Privacy characterisation of 4sq. behaviour

- check in – only shared with friends
  list of mayorship, tips, done – publically available.
- Result : possible to know home location of 78% of users with around error of 50 Kms.
- home city is not mandatory in 4sq.

Dataset :

- Tips can serve as feedback, recommendation or review to help others.
- Used Yahoo! Place finder API
  98% users provide valid location in home city
  13% users provide valid location in venue.

  ambiguous location users 2.7%
  venues 29.6%.

- around 30% of users in DB have atleast 1 done, mayor, venue
- few users have several (may. done, tip) while majority have only one (may, done, tip)
- high ~~co~~ correlation b/w no. of may. & no. of tips.
- high " " " & no of dones.
  (Spearman correlation)
- cities with largest no. of mayorship tend to have largest no. of tips and dones.

interesting things :

mayorship is more concerned in South-E asia
(Jakarta, Bandung, Singapore)

Tips are concentrated in several part of world
(N.Y, Jakarta, Sao paulo)

Done are concentrated in NY, Chicago, SF.

Temporal and spatial analysis.

- Temporal analysis.
  50% of users post consecutive tip/ done 1 hr apart
  50% of " have avg interactivity time = 950 hrs
  80% of users have max " time (167 hrs)

- Spatial analysis
  36% of users have max disp of 0 Km
  70% of users have a avg disp. of 150 Km

- return time
  69% users Returntime < 360 hrs

inferring home location

assumption: user tend to have mayorship, tip dones in venues at place where they live

Group users in 3 class. class 0: single activity
class 1: multiple acti with predominent location
class 2: multiple act. where no single position is great

Results:

1) best single attribute mayorship > tips > Dones. } for home location

2) best model All > ~~mayorship + Tip~~ } for home city (Tips only)

46% of distances are under 50 Km. predict
78% of users + predict + under 50 Km

for home country 90% acuracy tip > mayor > Dore

home state best - All mo

Research Ppr 2

name change in Twitter

- username can change, but user ID can't be changed
- Around 73-21% of user change attribute, 10% change username
- 20% of users trigger 85% of username change (Pareto principle)
- 65% of users choose a new username un-related to the old one while thirty five percent re-use an old one sometime later.
- Reasons to change: space gain, suit to trending event, lose / gain anon, adjust to real life event, avoid boredom, obscured username from, username squatting
- 82% of username is non-related to old username

① Popularity vs frequency of change

- may lead to confusion,

username frequency is weak yet positively correlated with in degree of user

② Activity vs frequency of change

change frequency is weak yet positively correlated

③ Familiarity vs frequency of change

both older & newer acc. engage in this behaviour.

Reasons in detail again - space gain - 140 char limit

75.19% people moved to short username
60.% people took longer name
~~words~~ ++ f(username) < 11 tend to add char

maintain multiple accn
anon → real → anon

major of ppl change to gain anon (30%)
violate wiki policy (22%)

- FB inspector outputs 2 values malicious / normal & confidence
- To check mongo has been successfull installed, type mongo
- MySQL login via terminal mysql -u (username) -p (db name)
  mysql -u username -p
- To fetch past tweets of Twitter user by using
  user_timeline endpoint of streaming API
- track → used to collect tweet mentioning a keyword (track tweet specific to geolocation)
- follow → used to collect tweet tweeted by user
  (track twee specific keyword)
- Prob of female acc & male req → will accept a female friend req.
- Latanya Sweeney's combined Zip, BD, Sex.
- Responsible for making deidentification difficult
  - pseudo names
  - multiple matches with similar attribute
  - lack of common attribute
- Graph API returns 25 results
- We cant get previous tweets by streaming API, only tweets by user we can get.

create db in mongodb → use (dbname)

FB1 → Ground truth extraction → generalising feature vector → Supervised learn. → Rest API

Feature belonging to Tweet metadata used in Tweet crawl

a) Source of tweet & Geocordinate.

- Token on FB
  User access
  Page access
  App access
- comp req to generate extended token - ~~App permission~~
  App ID, App secret, Short lived token.
- needed to create new app on fb → valid ph no, email, website
- search using graph API - User, page, ~~post~~, event, group, place
- Mongodb stores data in JSON format.
- following are measure of social & cognitive process of measure of Behavior
  Interpersonal focus, Social orientation, cognition, valence

females were more vulnerable to phishing attack

degree centrality: no. of ties a node has

most Reasons that rumour tweet is greater.
- ppl post rumour
- rumours are more believi

week 11

- evaluate the rank of tweets by Tweet cred NDCG
- fake images from real images tweet - Tweet based feats
- 70% whisper delete in week

week 10

- most accurate in infering home city - twitter
- median length 11.
- fake content via mobile device
- city from cn + data - educational model

- increasing order of accuracy of inferig user home city 4sq, Gt, Twitt, ~~fb~~ Zgeda

week 9

- check in is not publically available by default.
- check in are less useful than tip to reveal user hood

week 8.

- ppl change username for
  - squatting, space gain (indirect)
- density = edges / nodes
- betweenness centrality: measure of centrality of graph based on shortest path. for every pair of vertices in connected graph there exist atleast one shortest path

week 7

10 mails were spoofed by (0 email id of person did not know

week 6

PageRank is a algorithm used by G to rank website.

Klout score : social media influence.

Cronbach alpha; internal consistency, how closely related a set of items as a group

Cohen's kappa: measure inter rater agreement for categorical item

social influence: no of follow
social reputation : like, love, rating

week 2

- ~~API~~ retrieve data in JSON & XML

- classification of fake image
  - Tweet based feature performed better than user based feature

- Decision Tree better than Naive Bayes

→ punkt
→ from collection import
Port Stemmer
Stop words | ngrams

Twecon - command line Twitter data collection tool

python ~~type~~ ~~tqdm~~ library pymongo, MySQL-python (mysqldb)