Chapter 19
Social Media Sharing

19.1 Representative Social Media Services

- Social media, a group of Internet-based applications based Web 2.0, allow the creation and exchange of user-generated contents.
- Key factors to the success of the new generation of social media:
  - Collective intelligence
  - Rich connections and activities
- Two important social media services:
  - User-generated content sharing
  - Online social networking

User-Generated Content Sharing

- *UGC* is used for a wide range of applications with different types of media (e.g., *text, music, picture, and video*).
  - Video data are arguably more difficult for content generation and sharing, given their large size, high bandwidth, and long playback duration.
- **YouTube**: the most significant and successful video sharing website
  - Fast growth (established in 2005):
    - In 2006: 100 million videos were served per day
    - By 2020:
      - Over 2 billion unique users visit YouTube per month
      - 500 hours of new videos uploaded every minute
      - Over 70% views are on mobile
  - Highly globalized:
    - 91 countries and across 80 languages
    - 85% of YouTube traffic comes from outside of the US

Online Social Networking

- *OSN* provides an Internet-based platform to connect people with social relations.
- Both **Facebook** and **Twitter** support the sharing and propagation of such media objects as *pictures, music, and video* among friends.
  - Facebook:
    - 2.5 billion active users.
    - 1.66 billion million of them log on a daily basis.
  - Twitter:
    - A representative of microblog, recently has also begun offering the Vine service, which enables mobile users to create and post video clips up to six seconds long.
    - 330 million monthly active users.
    - 145 million daily active users.

19.2 User-Generated Media Content Sharing

- Understanding the features of social media services is crucial
- Using YouTube as a representative*:
  - YouTube video Format and Meta-data
  - Characteristics of YouTube Video
  - Small-World in YouTube Videos
  - YouTube from a Partner's View
  - Enhancing UGC Video Sharing

* The YouTube dataset is from X. Cheng, J. Liu, C. Dale, "Understanding the characteristics of internet short video sharing: a YouTube-based measurement study".

YouTube Video Format and Meta-data

- **Format**:
  - Uploaded videos in many formats are converted to the .FLV (Adobe Flash Video) format after uploading.
  - The H.263 video codec was used earlier, the H.264 video codec was introduced in late 2008 for "high quality".
  - Later added open VP9 codec, 4K/8K resolution, and 360 degree video.

YouTube Video Format and Meta-data (Cont'd)

- **Meta-data**:
  - video ID (an 11-digit ID composed of the characters 0–9, a–z, A–Z, -, and _), uploader, date added, category, length, number of views, number of ratings, number of comments, and a list of related videos.

ID	YiQu4gpoa6k
Uploader	NewAgeEnlightenment
Date added	August 08, 2008
Category	Sports
Video length	270 s
Number of views	924,691
Number of ratings	1,039
Number of comments	212
Related videos	ri1h2_jrVjU, 0JdQlaQpOuU, …

Table 19.1: An example of the meta-data of a YouTube video

Characteristics of YouTube Video

- **Video Category**
  - 15 categories
  - Highly skewed distribution: "Entertainment" and "Music" videos constitute half of the entire collection

Rank	Category	Count	Percentage (%)
1	Entertainment	1,304,724	25.4
2	Music	1,274,825	24.8
3	Comedy	449,652	8.7
4	People and blogs	447,581	8.7
5	Film and animation	442,109	8.6
6	Sports	390, 619	7.6
7	News and politics	186,753	3.6
8	Autos and vehicles	169,883	3.3
9	Howto and style	124,885	2.4
10	Pets and animals	86,444	1.7
11	Travel and events	82,068	1.6
12	Education	54,133	1.1
13	Science and echnology	50,925	1.0
14	Unavailable	42,928	0.8
15	Nonprofits and activism	16,925	0.3
16	Gaming	10,182	0.2
17	Removed	9,131	0.2

Table 19.2: List of YouTube video categories

Characteristics of YouTube Video (Cont'd)

- **Video Length**
  - The most distinguishing difference from traditional video contents: YouTube mostly comprises short video clips

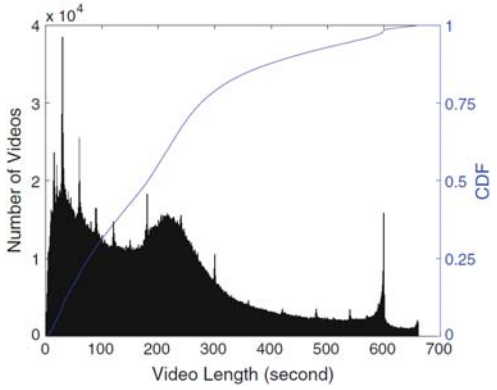


Fig. 19.1 Histogram and cumulative distribution (CDF, the solid line) of YouTube video length

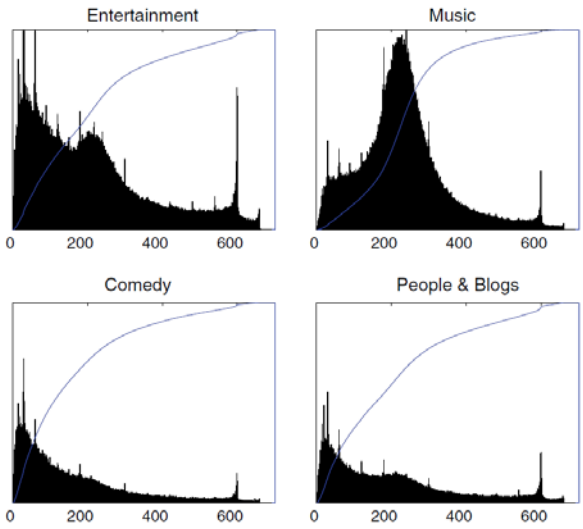


Fig. 19.2: Length histograms and cumulative distributions for the four top categories

Characteristics of YouTube Video (Cont'd)

- **Access Patterns**
  - 10% top popular videos account for nearly 80% of views
  - 60% of videos are watched for less than 20% of their duration
  - Only 10% of the videos are watched again on the following day
  - Copyrighted videos tend to get most of the views earlier
  - Top videos tend to experience sudden bursts of popularity
  - Users' viewing behaviors are affected by both the video quality and their social relations

Small-World in YouTube Videos

- In YouTube, there is a strong correlation between the number of views of a video and that of its top related videos.
  - The graph topology for the network of YouTube videos can be measured by using the related links to form directed edges in a video graph.

- A small-world network has a large clustering coefficient like a regular graph, but it also has a small characteristic path length like a random graph.
- The clustering coefficient of a node is the ratio of existing links connecting a node's neighbors to each other to the maximum possible number of such links. The clustering coefficient for the entire network is the average of the clustering coefficients of all the nodes.
- The characteristic path length is calculated by finding the shortest path between all pairs of nodes, adding them up, and then dividing by the total number of pairs.
- A random graph is generated based on certain probability distributions. Purely random graphs, built according to the **Erdös-Rényi** (**ER**) model, exhibit a small characteristic path length (varying typically as the logarithm of the number of nodes) along with a small clustering coefficient.
- A regular graph is a graph where each vertex has the same degree.

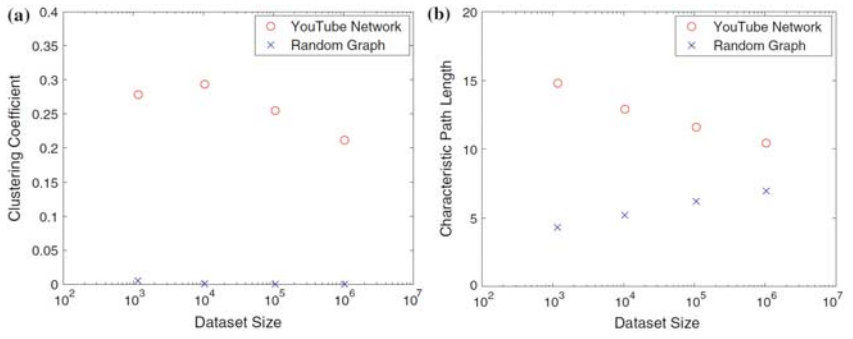


Fig. 19.3: Small-world characteristic of YouTube videos

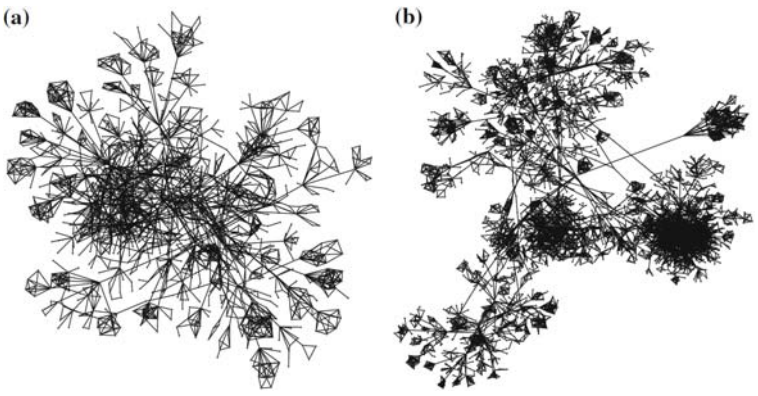


Fig. 19.4: Two sample graphs of YouTube videos and their links

YouTube from a Partner's View

- The main source of YouTube's revenue is from advertisements.
- A *YouTube Partner Program* was introduced to accommodate the content owners with copyrighted videos and popular channels.
- The *Insight Analytics* from YouTube can provide various basic statistics on videos and channels so that the YouTube partners can leverage these statistics to adapt their content deployment and user engagement strategies.

YouTube from a Partner's View (Cont'd)

- The *referral sources* (the last webpages where the viewers come from) can be classified into five categories.
  - Suggestion: YouTube's related video links
  - Video Search: YouTube or search results, e.g., from Google
  - YouTube Surfing: YouTube pages except related video links and search results
  - Social Referral: a link on an external webpage, or the video was embedded on an external webpage
  - Non-Social Direct: the viewer navigated directly to the video

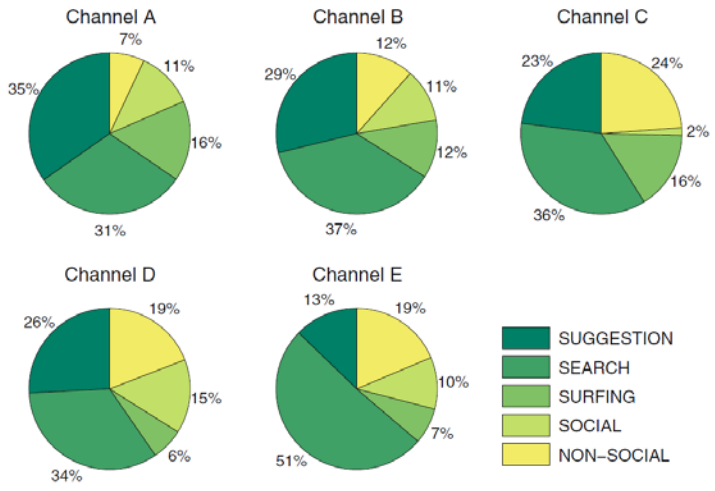


Fig. 19.6: Breakdown of the referral source

YouTube from a Partner's View (Cont'd)

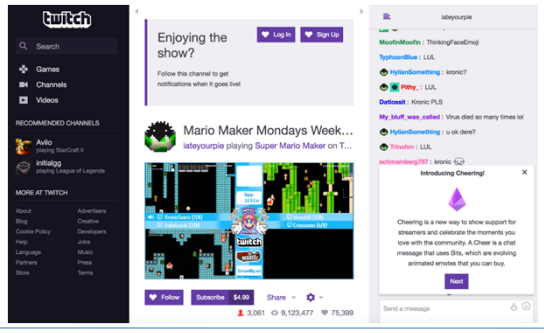
- Search results (Video Search) and related videos (Suggestion) are the top sources of views.
- Although Social Referral is not the top view source, the impact of external website referral cannot simply be ignored.

Table 19.3: Summary of top external websites referrers

	Channel A	Channel B	Channel C
1st	9.0 % downloading site	16.2 % Facebook	31.9 % gaming wiki
2nd	4.4 % Facebook	2.2 % n/a	7.6 % Facebook
3rd	2.6 % forum	1.5 % downloading site	5.3 % gaming blog
4th	1.7 % gaming site	1.2 % n/a	5.1 % gaming site
5th	1.5 % gaming site	0.9 % downloading site	3.7 % Internet video site
	Channel D	Channel E	
1st	41.2 % Reddit	62.4 % Facebook	
2nd	9.9 % Facebook	2.4 % music streaming	
3rd	4.7 % Twitter	2.0 % music blog	
4th	2.0 % blog	2.0 % Twitter	
5th	1.7 % entertainment site	1.6 % music blog	

Crowdsourced Interactive Livecast

- Numerous amateur broadcasters stream their own contents to viewers in their channels
  - Examples: Twitch.tv3 , Youtube Gaming4, Mixer5 , Inke.tv6, etc.
  - A viewer also interacts with the broadcaster and other viewers within the same channel.



Viewer Interactions

- The viewers of FIFA channels are more active with frequent interactions than GS:GO.

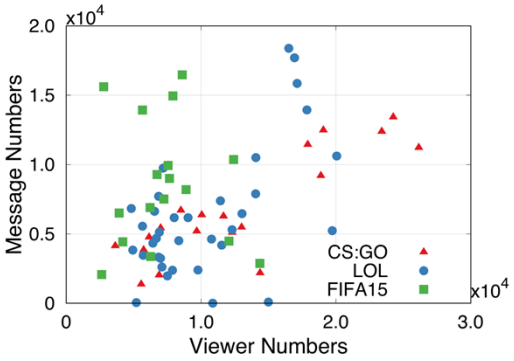


Fig. 19.8: The distribution of viewer numbers and interaction message numbers of channels in three games.

Viewing and Interaction Patterns

- Watching preference vary with different time periods and/or different channels

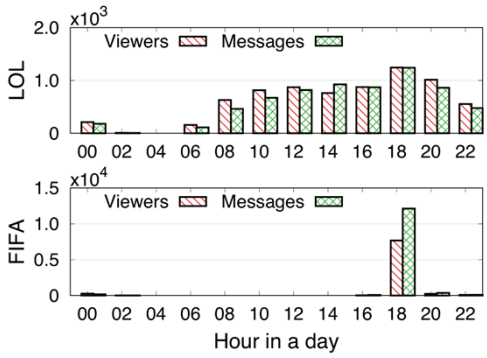


Fig. 19.9: The average viewing and interaction numbers of different time periods in a day for two gamecast channels.

Viewer Interactions

- 87% sessions never generating messages and more than 5% sessions having frequent messages (more than 10 messages per viewing).
  - Viewers with frequent interactions generally prefer low streaming delay.
  - Silent viewers with less involvement are not so sensitive to delay

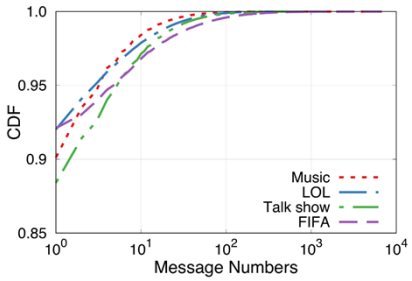


Fig. 19.10: The CDF plot of interactive message numbers of each session in different channels.

19.3 Media Propagation in Online Social Networks

- Video has become an important type of object spreading over social networks, beyond earlier simple text or image object sharing.
- Video sharing services and social networking services have become highly integrated.
  - 500 years of YouTube video are watched every day on Facebook
  - Over 700 YouTube videos are shared on Twitter each minute
- Significant challenges are presented not only to social networking service management, but also to network traffic engineering and to the resource provisioning of external video sites.

Sharing Patterns of Individual Users

- Video object sharing involves both propagation over the online social network and accesses to the external video site
  - How often do users initiate video sharing?
  - How often do users further share a video upon receiving it?
- The used dataset*:
  - one week long, 12.8 million video sharing events, 115 million viewing events, 827 thousand initiating records
- Three types of users according to activeness:
  - *Spreaders* (SU)
  - *Free-riders* (FU)
  - *Ordinary users* (OU)

* The dataset is from X. Cheng, H. Li, J. Liu, "Video sharing propagation in social networks: measurement, modeling, and analysis".

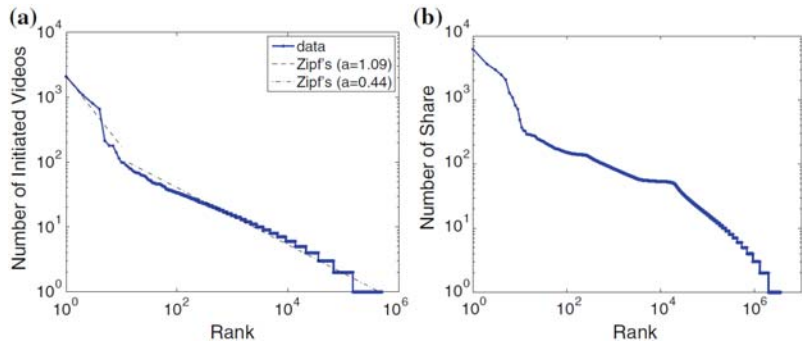


Fig. 19.11: Rank distributions of **(a)** Initiated videos **(b)** Shared videos

Video Propagation Structure and Model

- Two typical propagation structures:
  - a moderate depth, but limited branching
  - frequently branching at different levels

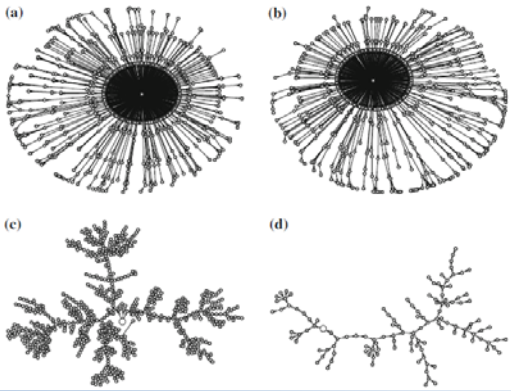


Fig. 19.12: Illustration of spreading trees for popular videos.
(a) size = 1093, height = 9
(b) size = 951, height = 8
(c) size = 805, height = 30
(d) size = 126, height = 23

Video Propagation Structure and Model (Cont'd)

- **The SIR model (Susceptible-Infectious-Recovered):**
  - At a particular time t,
    - $S(t)$ – The number of individuals not yet infected;
    - $I(t)$ – The number of individuals who have been infected and are capable of spreading the disease to those in the susceptible category;
    - $R(t)$ – The number of individuals who have been infected and then recovered.
- Given transition rate β from S to I and γ from I to R

$$\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I$$

- There is a natural mapping for object sharing propagation:
  - All the users in the social networks are Susceptible at the beginning; at a certain time, the users accessing the object are Infectious, indicating that they are able to infect others by sharing the object; they can be Recovered if they choose not to share.

Video Propagation Structure and Model (Cont'd)

- For video spreading, the mapping is not complete:
  - **Immune (Im):** indicating users who choose not to watch the received video, and likely not participate in the spreading as well.
  - **D1 and D2:** two temporary decision stage (users decisions on watch or not, and share or not share)
  - **Permanent (P):** indicating users who have shared the video after watching the video to differentiate the users that are Recovered (Watch, but not share)

- With the enhanced SI^2RP, the probability distributions that characterize the behavior of different types of users (SU, FU, OU) can be inferred from measurement results.

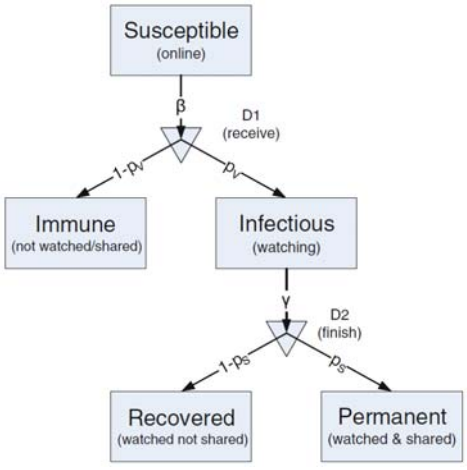


Fig. 19.13: The Susceptible-Immune-Infectious-Recovered-Permanent (SI^2RP model)

Video Watching and Sharing Behaviors

- Users are more likely to browse the contents from friends, and are willing to share their resources to assist friends.
- Three options upon receiving the video post:
  - Watches the live video (*streaming users*);
  - Download the video and expect to watch it later (*storage users*);
  - Shows no interest in the video.
- Streaming users
  - Streaming quality is required (e.g., startup latency and playback continuity)
  - Dynamic – users may lose interest and stop watching
- Storage users
  - No such concerns as interest nor playback quality until watching the video
  - Relatively stable – users download the video asynchronously.

19.4 Mobile Video Clip Sharing

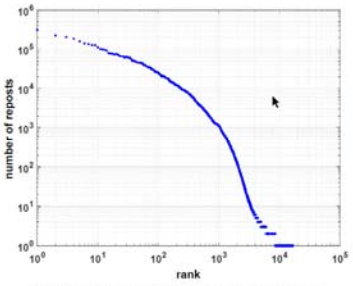
- New generation video sharing services that use smart mobile terminals to instantly capture and share ultra-short video clips (usually of several seconds).
- Representatives: Twitter's Vine, Instagram, Snapchat, TikTok.
  - TikTok (known as Douyin in China) has become the 8th most-downloaded app in the past decade.

19.4.1 Mobile Interface Characteristics

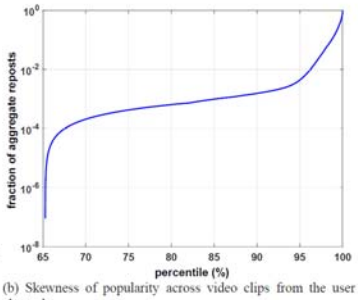
- An instant video clip itself is of only several seconds long, thereby a mobile user can hardly tolerate a long delay,.
- Solution: Return a playlist of video clips when a user touches the screen to view the updates for certain users, tags, or channels.
- *Batched view* with scrolling
  - A series of user gestures, e.g., click, drag and fling
  - Speed, acceleration, and continuity vary depending on user's input.

19.4.2 Video Clip Popularity

- popularity of video clips is extremely skewed
  - top 5% video clips accounts for more than 99% reposts.
  - heavily deviates from the Pareto Principle (or 80-20 rule).



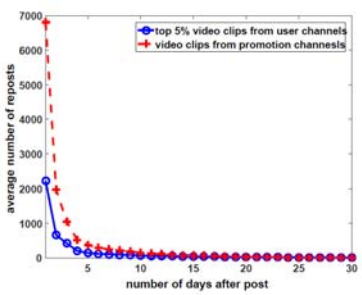
(a) Video clips rank ordered by the number of repost



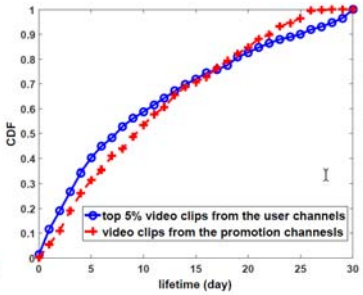
(b) Skewness of popularity across video clips from the user channels

19.4.3 Lifespan and Propagation

- More than half of the popular video reposts only stay in active for less than 10 days.
  - quite different from traditional video sharing services.
  - some of YouTube videos can still get views even after one



(a) Daily number of reposts



(b) CDF of lifespan

19.5 Further Exploration

- Research on the social relations and social graphs in the human society has a long history, so does that on disease propagation in epidemiology.

- Online social media and social networking however appeared only in very recent years and are still undergoing rapid changes. Research in this field remains in an early stage and many exciting topics are to be explored.