Content Viralization

- Web 2.0 and User Generated Content
- Understanding Viral Content
- Paths to Viralization
- Youtube, Instagram and Tiktok's Recommendation System
- Network Effect
- Sequential Decision Model
- Information Cascade

Web 2.0 Platforms

Web 2.0 Platforms

Web 1.0 (1990 - early 2000s)

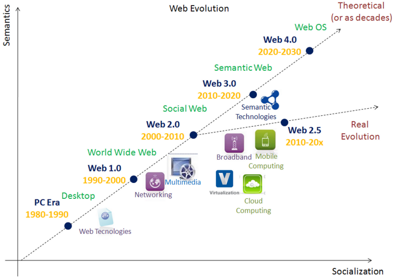
- aka “read-only web”
- HTML-based, static, informational, non-interactive websites
- basic content– text, images, and hyperlinks

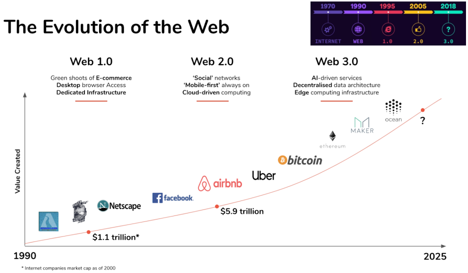
Web 2.0 (2000s - Present)

- aka “read-write web”, “social web”
- Dynamic user-generated content
- Social media, e-commerce, blogs
- Real-time interactivity
- personalization possible

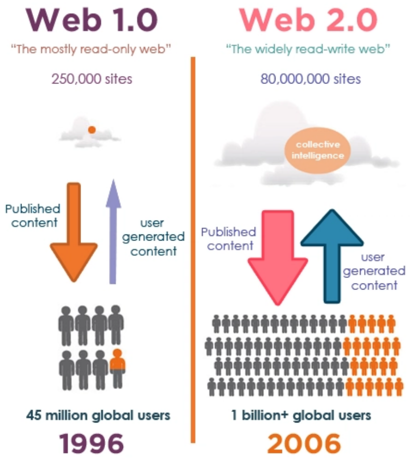
Web 3.0 (Future)

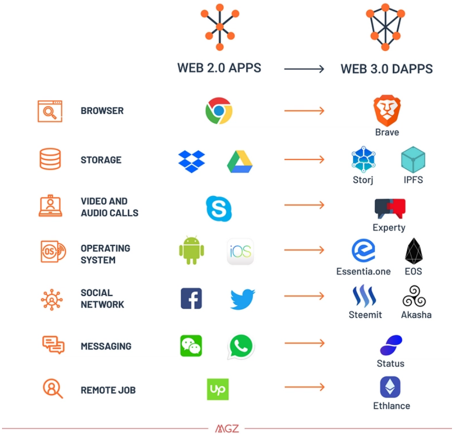
- “read-write-execute”
- “semantic web” or “decentralized web”
- machine-to-machine interaction
- AI, ML, blockchain
- Personalized, privacy-focused





Web 2.0 Platforms

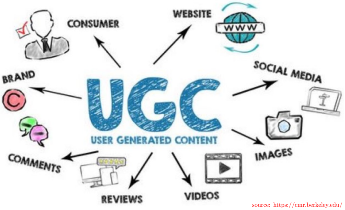


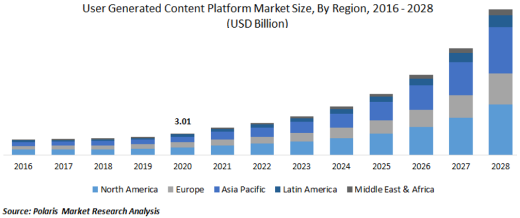


source: https://dev.to/pragativerma18/evolution-of-web-42eh

User-Generated Content (UGC) and Its Impact on Social Media

UGC powers Web 2.0, turning users into content creators





- Content Creation: Videos, posts, blogs (e.g., YouTube, Instagram)
- Engagement: Likes, shares, comments fuel platform growth
- Network Effects: Content spreads quickly via shares, amplifying reach

Impact on Platforms:

- Diverse Content: Expands user engagement across demographics
- Algorithmic Promotion: Recommends content based on interactions
- Influencer Culture: Creators drive trends, monetization

The Rise of Viral Content in Web 2.0 Era

Viral content is a common phenomenon due to participatory Web 2.0

- Ease of Sharing: Social networks amplify content reach (e.g., Facebook, Twitter)
- Algorithmic recommendations: Platforms promote viral content based on engagement metrics
- Tech Access: Mobile devices, apps like video editors make content creation accessible
- Engagement Drives Virality: Likes, shares, comments fuel exponential growth
- Network effects: Viral content spreads exponentially as more people engage with it

The Role of Viral Content in Social Media Growth

Viral content drives platform growth and user engagement:

- Platform Growth: Viral content attracts new users and boosts activity
- Monetization: Generates ad revenue and drives subscription models (e.g., YouTube, TikTok)
- Brand Visibility: Businesses leverage viral content for low-cost marketing

Impact on Metrics:

- User Acquisition: Trending videos, memes, and challenges attract new users
- Increased Engagement: Higher watch time and user interaction
- Revenue Growth: Viral content boosts ad revenue and premium subscriptions

YouTube's monetization through viral content and creators' ad revenue

Viral Content and Community Building

Viral content fosters deep user engagement and creates communities:

- Shared Experiences: Memes, trends, and challenges unite users
- Comment Sections: Forums and discussions spark community building
- User Participation: Challenges and trends encourage user-generated content (e.g., TikTok)
- Platform Algorithms: Reward engaging content with visibility, fueling more creation
- Global Communities: Viral phenomena like the "Ice Bucket Challenge" raised awareness for ALS and build worldwide support

Paths to Viral Content

Key Paths to Viral Content

Content can go viral through various paths, driven by user behavior, algorithms, and network dynamics Key mechanisms include:

- Web Search: Content discoverability through search engines (e.g., Google)

- Subscriptions: Regular content delivery that fosters repeat engagement

- Referrals: Sharing via email, social media, or word-of-mouth

- Recommendations: Algorithms suggest content to users based on behavior (e.g., YouTube, TikTok)

Leveraging Web Search & SEO

Users discover new content through Web search

Optimizing content for search engines can increase its discoverability:

- SEO: Optimizing content to rank higher in search engine results
  - Keywords: Identifying and incorporating the right keywords
  - Backlinks: Gaining links from reputable sites
  - On-page optimization: Proper use of headings, meta descriptions, and alt text for images
  - User experience: Fast page load times, mobile-friendliness, easy navigation
- Relevance and Keywords: maximize chances to match user queries
- Content Structure: Proper formatting, titles, and meta descriptions improve search indexing

Blog posts or YouTube videos with optimized titles and descriptions appear at the top of search results

Subscriptions and Referrals for Viral Spread

Subscriptions: Strengthen viral potential by building a loyal subscriber base, ensuring strong early engagement.

- Subscription Engagement:
  - Notifications boost immediate visibility and engagement
  - Consistent posting schedules foster routine interaction
  - Exclusive or early-access content rewards loyal subscribers
  - Calls to action encourage sharing, likes, and comments
  - Direct engagement (comments, replies) nurtures a community

Referrals: Amplify reach through sharing, social media, and word-of-mouth

- Leveraging Referrals:
  - Incentivize referrals (e.g., discounts, exclusive access)
  - Harness social trust: recommendations from friends are powerful
  - Monitor referral engagement to optimize viral strategies

Recommendations & Algorithms Driving Virality

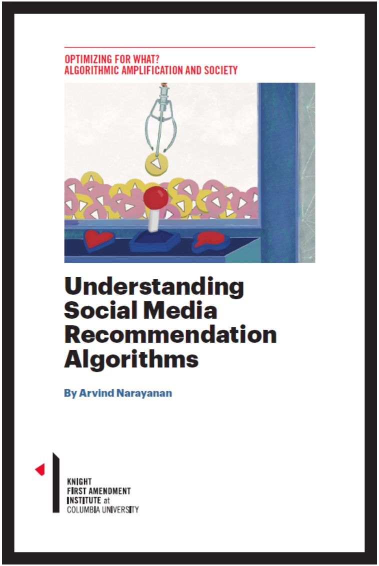
Algorithms help viralize content, shaping what users see and engage with:

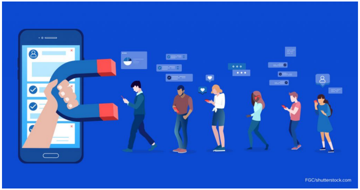
- Personalized Discovery: Algorithms analyze user behavior (watch time, likes, shares) to recommend relevant content
- Feedback Loops: As content gains engagement, it is further promoted, creating exponential growth

Examples of Algorithmic Recommendations:

- Home page recommendations: YouTube curates videos based on individual watch history
- Related videos: Suggests similar content to keep users engaged
- Trending section: Highlights popular videos based on real-time engagement metrics
- Shorts boosting full videos: YouTube Shorts often link users back to full-length content, driving viral loops
- TikTok's "For You" page: Surfaces highly engaging short videos to massive audiences

Social Media Recommendation Algorithms





Algorithmic Amplification of Extremism

From clicks to chaos: How social media algorithms amplify extremism

AUTHOR : SOUMYA AWASTHI

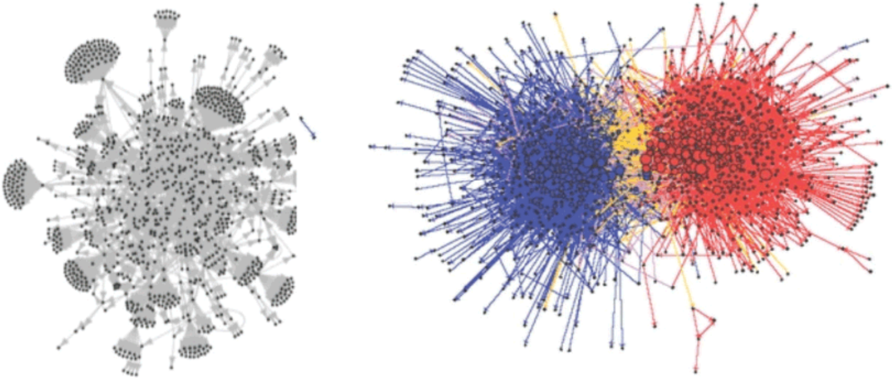
ORF OBSERVER RESEARCH FOUNDATION

Ideas · Forums · Leadership · Impact

- Social media algorithms prioritize content that generates high user engagement
- This prioritization often elevates emotionally charged or sensational content, including extremist material
- This can lead users into ideological echo chambers, filter bubbles, and polarization
- Reinforcing existing beliefs can potentially radicalizing users
- Personalized content delivery limits exposure to diverse views, deepening ideological divides
- Extremist content spreads, inciting violence, misinformation, and eroding trust in institutions, destabilizing society
- Extremist groups exploit these algorithms to spread propaganda and recruit followers across platforms like YouTube, TikTok, and X (formerly Twitter)
- The phenomenon of 'algorithmic radicalization' demonstrates how users can be guided toward progressively more extreme content
- Addressing this issue requires a combination of technological solutions, regulatory measures, and public awareness initiatives

Polarization in Social Network

source: Ulicny, Kokar, Matheus, (2010)



(a) Malaysian Sopo blogosphere (b) US political blogosphere (Adamic & Glance, 2005[4])

A visualization of Malysia and US blogosphere (nodes are blogs and edges are links to blogs). Left reveals importance/credibility/popularity of blogs, while the right visual clearly show that bloggers are more likely to link to bloggers with the same party affiliations, forming two dense clusters with little interaction with the other cluster

Social Learning Disrupted by Algorithms

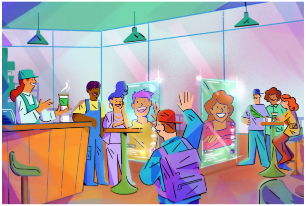
Social-Media Algorithms Have Hijacked "Social Learning"

BASED ON THE RESEARCH OF

William Brady
Joshua Conrad Jackson
Björn Lindström
M.J. Crockett

KelloggInsight

A magazine of research and ideas from the faculty of the Kellogg School of Management



Jesús Escudero

- Social learning—learning by observing others—is fundamental to human behavior
- Social media algorithms disrupt this process by filtering content based on engagement, not relevance
- This leads to exposure to emotionally charged, sensational, or extreme content, skewing perceptions
- Algorithms prioritize content that is prestigious, in-group, moral, or emotional (PRIME), which can distort understanding
- Adjusting algorithms to promote diverse, accurate content can restore healthier social learning

Social Media Algorithms Amplify Misogynistic Content to Teens

Social media algorithms amplify misogynistic content to teens

5 February 2024

Social media algorithms amplify extreme content, such as misogynistic posts, which normalises harmful ideologies for young people, finds a new report led by a UCL researcher.

- A UCL-led study found a fourfold increase in misogynistic content on TikTok's "For You" page over five days
- Algorithms promote extreme content by targeting user vulnerabilities like loneliness and loss of control
- Misogynistic tropes have moved from screens into schools, normalizing harmful ideologies among youth
- Recommendations include a "healthy digital diet" education, peer-to-peer mentoring, and involving boys in discussions on online misogyny
- The study calls for social media companies to prioritize user well-being over profit

Weaponisation of Social Media Algorithms

Would I Lie to You? The Weaponisation of Social Media

DECEMBER 5, 2022

Global

Policy Centre for Governance

- Social media algorithms, designed to maximize user engagement, often amplify divisive and harmful content
- This amplification can distort public discourse, erode trust in democratic institutions,and incite violence
- Vulnerable groups, including ethnic and religious minorities, are particularly susceptible to targeted disinformation campaigns
- Combating this issue requires a multifaceted approach, including regulatory measures, technological solutions, and public awareness initiatives
- UNDP is actively involved in promoting information integrity through fact-checking, digital literacy programs, and supporting democratic processes

Weaponising Social Media for Information Divide and Warfare

Weaponising Social Media for Information Divide and Warfare

Authors: Ehsan-Ul Haq, Gareth Tyson, Tristan Braud, Pan Hui | Authors Info & Claims

HT '22: Proceedings of the 33rd ACM Conference on Hypertext and Social Media • Pages 259 - 262
https://doi.org/10.1145/3511095.3536372

Published: 28 June 2022 Publication History

Check for updates

- Social media platforms play a dual role during crises: they can both alleviate and exacerbate challenges
- Information polarisation on these platforms can deepen societal divides
- Malicious actors exploit these divides, using social media for information warfare
- COVID-19 Pandemic: Social media platforms were inundated with misinformation and conspiracy theories about the virus and vaccines. Algorithms amplified such content due to high engagement metrics, leading to public confusion and undermining trust in health guidelines
- Russia-Ukraine Conflict: Both state and non-state actors utilized social media to disseminate propaganda, disinformation, and manipulate public opinion. Platforms like Facebook and Twitter became battlegrounds for information warfare, with limited content moderation allowing the spread of misleading narratives
- Addressing these issues requires a multifaceted approach, including policy interventions, technological solutions, and public awareness

Algorithmic amplification of politics on Twitter

Ferenc Huszár[a,b,c,1,2], Sofia Ira Ktena[a,1,3], Conor O'Brien[a,1], Luca Belli[a,2], Andrew Schlaikjer[a], and Moritz Hardt[d]

[a]Machine Learning Ethics, Transparency, and Accountability Team, Twitter, San Francisco, CA 94103; [b]Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, United Kingdom; [c]Gatsby Computational Neuroscience Unit, University College London, London, W1T 4JG, United Kingdom; and [d]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720

Twitter's algorithms tend to amplify political content, making it more visible to users

- Polarization: This can contribute to increased political polarization by exposing users predominantly to content that aligns with their existing beliefs
- Echo Chambers: Users may find themselves in echo chambers, where their views are reinforced, and opposing perspectives are underrepresented
- Algorithmic Bias: May unintentionally favor sensational or divisive content, further entrenching divisions
- Public Discourse: Can significantly influence the nature of public discourse, potentially skewing it towards more extreme viewpoints
- Democratic Processes: Can affect democratic processes by shaping public opinion and discourse in ways that may not reflect a balanced representation of views

- Algorithm Transparency: Advocate for greater transparency in how algorithms prioritize and recommend content
- Diversification of Content: Encourage platforms to diversify the content presented to users to expose them to a broader range of perspectives
- User Awareness: Promote user awareness about the influence of algorithms on the content they see and engage with

Algorithmic Amplification of Politics on Twitter

Governments worldwide are implementing measures to mitigate the spread of harmful content:

- Brazil: Temporarily banned X until it appointed a legal representative and blocked accounts questioning election legitimacy.
- European Union: Introduced rules threatening fines and suspension for platforms failing to prevent election interference.
- United Kingdom: Enacted an online safety act compelling social media sites to tighten content moderation.
- United States: Proposed legislation to ban TikTok unless sold by its Chinese parent company.

Free Speech Concerns: Critics argue that such regulations may infringe upon free speech and internet principles, advocating for a marketplace of ideas without governmental interference

Pros and Cons of Social Media Algorithms

- Enhanced Content Discovery: Algorithms help users find valuable resources and connect with like-minded peers, especially for marginalized groups
- Efficient Content Moderation: Automated systems detect and remove harmful content, including self-harm and cyberbullying, fostering a safer online environment
- Support for Political Engagement: Algorithms can connect users with like-minded individuals, facilitating collective action and raising awareness about societal issues
- Real-Time Information Sharing: Platforms use algorithms to attach accurate voting information to posts about elections, promoting civic engagement
- Improved User Experience: Personalized content recommendations enhance user satisfaction by tailoring feeds to individual preferences

- Mental Health Risks: Prolonged engagement driven by algorithms can lead to exposure to age-inappropriate content, unrealistic beauty standards, and cyberbullying, adversely affecting youth mental health
- Privacy Concerns: Algorithms track and exploit users' online behavior, raising issues about data collection and targeted advertising without explicit consent
- Echo Chambers and Polarization: Personalized feeds can create filter bubbles, limiting exposure to diverse viewpoints and reinforcing existing beliefs, potentially leading to political polarization
- Algorithmic Bias: Unconscious biases in algorithm design can perpetuate stereotypes and disproportionately affect marginalized communities
- Lack of Transparency: The "black box" nature of algorithms makes it difficult for users to understand how content is curated, leading to confusion and mistrust

Psychological Impacts of Algorithmic and AI-Driven Social Media on Teens

The Psychological Impacts of Algorithmic and AI-Driven Social Media on Teenagers: A Call to Action

Sunil Arora, Sahil Arora, John D. Hastings

- Mental Health Concerns: Prolonged social media use is linked to anxiety, depression, and other psychological issues among teenagers
- Addiction and Screen Time: Continuous scrolling and notifications contribute to excessive screen time, affecting sleep patterns and overall well-being
- Diminished Attention Span: The constant influx of information leads to reduced attention spans and challenges in focusing on tasks
- Sleep Disruption: Engaging with social media late at night disrupts sleep, impacting cognitive functions and mood
- Algorithmic Manipulation: Platforms use algorithms to maximize engagement, often at the expense of user well-being
- Lack of Transparency: Users are often unaware of how algorithms curate content, leading to concerns about privacy and control
- Exposure to Harmful Content: Algorithms may inadvertently promote content that is harmful or inappropriate for teenagers
- Peer Pressure and Comparison: The curated nature of social media fosters unrealistic comparisons and peer pressure

Psychological Impacts of Algorithmic and AI-Driven Social Media on Teens

American Journal of Law & Medicine

ALGORITHMS, ADDICTION, AND ADOLESCENT MENTAL HEALTH: An Interdisciplinary Study to Inform State-level Policy Action to Protect Youth from the Dangers of Social Media

Published online by Cambridge University Press: **12 February 2024**

Nancy Costello, Rebecca Sutton, Madeline Jones, Mackenzie Almassian, Amanda Raffoul, Oluwadunni Ojumu, Meg Salvia, Monique Santoso, Jill R. Kavanaugh and S. Bryn Austin

Show author details

Mental Health Issues:

- Exposure to harmful content leads to increased depression, anxiety, and body image concerns among adolescents
- Increased suicidality, eating disorders, and self-harm linked to algorithmic recommendations
- Vulnerability of adolescents' brain development increases emotional sensitivity and response to social media feedback

Addiction and Exploitation:

- Social media platforms use addictive design elements like endless scrolling and notification triggers
- Algorithms promote content that retains users, maximizing engagement and monetization
- Platforms are financially incentivized to keep youth engaged, profiting from targeted advertising

Algorithmic Amplification for Collective Intelligence

ESSAYS AND SCHOLARSHIP

Algorithmic Amplification for Collective Intelligence

Social media promised a new, democratized, and digital public sphere. Algorithms can help us get there.

BY JASON W. BURTON
SEPTEMBER 21, 2023

KNIGHT FIRST AMENDMENT INSTITUTE at COLUMBIA UNIVERSITY

From “Liberation Technology” to the “Post-Truth Era”

- Social media platforms as democratizing tools for public discourse
  - Algorithms can enhance the reach of diverse voices and perspectives
  - Properly designed algorithms can support informed and reflective deliberation
- Algorithmic curation can limit achieving a truly informed public sphere
  - Engagement-based ranking prioritizes sensational content, concerns over misinformation and polarization
  - This can undermine informed deliberation and democratic engagement
- Goal: align algorithmic curation with democratic values
- Shift from engagement metrics to measures of deliberative quality
- Incorporate values like diversity, accuracy, and inclusivity into algorithmic design
- Examples include platforms like Polis that facilitate structured deliberation
- Balance algorithmic optimization with user autonomy

The Psychology of Viral Content

The Psychology of Viral Content

The psychology behind why certain content goes viral is closely tied to human motivations and emotional responses

Content that resonates with these emotions is more likely to be shared across social networks

Viral content often appeals to middle-stage needs in Maslow's hierarchy, such as belonging, self-identity, and self-fulfillment

- Viral content speaks to deep, universal human desires
- Emotional engagement is central to virality.
- Positive content performs better than negative content.

Motivations for Sharing Content

People share content for several reasons, often influenced by their emotional state and desire to connect:

- To inform others: Sharing valuable or entertaining content to enhance others' lives
- To define themselves: Sharing content that aligns with personal values or interests
- To connect with others: Strengthening relationships and maintaining social bonds
- For self-fulfillment: Gaining recognition and validation from peers
- To promote causes: Raising awareness for social or personal causes

These motivations align with Maslow's hierarchy, particularly focusing on esteem and belonging needs

Types of Content that Go Viral

Thematic content that engage viewers emotionally, encourage sharing, and typically goes viral taps into six key psychological drivers:

- Taboo: Content that challenges societal norms or is perceived as edgy
- Unusual: Unexpected or novel content that surprises and captures attention
- Outrageous: Content that shocks or pushes the boundaries of social acceptability
- Hilarious: Content that entertains, often invoking laughter or joy
- Remarkable: Content that is extraordinary, noteworthy, or highly creative
- Secrets: Content revealing something hidden/exclusive, provide a sense of privilege

source: GeeksforGeeks.com

The Role of Emotions in Viral Content

Emotions compel people to act and share content, often driven by a need for validation or connection. The following emotions that drive sharing:

- Awe: Content inspiring wonder or admiration (extraordinary achievements)
- Anger: Content that provokes outrage or disbelief (e.g., injustice or inequality)
- Surprise: Content that shocks or breaks expectations, evoking curiosity
- Joy: Content that lifts the mood, creates happiness, or offers entertainment
- Anxiety/Fear: Content triggering concern about loss/urgency (limited time offer)
- Lust: Content appealing to desires, related to love, success, or material gains

8 Basic Emotions and the purpose of each one

source: Calmsage.com

Psychological Mechanisms Behind Viral Content

Several psychological mechanisms explain why content becomes viral and help marketers and content creators maximize shareability

- Social Proof: The more people share content, the more others engage with it
- FOMO (Fear of Missing Out): Content that makes viewers feel they might be missing out on important or entertaining experiences
- Reciprocity: People share content to give back or feel sense of gratitude/obligation
- Self-Expression: Sharing content that reflects one's identity, values, or beliefs

Buzzfeed and Viral Content: A Case Study

Buzzfeed's Success with viral content is a result of understanding the psychology of sharing:

- Jonah Peretti, founder of BuzzFeed, emphasizes creating content that taps into universal, emotional responses
- BuzzFeed's content often evokes humor, surprise, or curiosity, making it highly shareable
- Social Vertices: The people who are connected to diverse social groups are key players in the viral process, acting as bridges between different networks

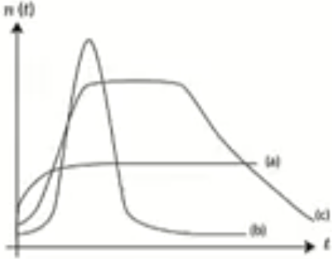
BuzzFeed's strategic use of emotional engagement and social connections is a proven model for viral content

Structural Virality and its Core Drivers

The Three Pillars of Virality

Viral content refers to the following three characteristics:

1. Height: Peak Performance
   - Massive viewership/adoption peaks
   - Breaks through content noise
   - Reaches significant thresholds
2. Longevity: Sustained Impact
   - Maintains momentum post-spike
   - Accumulates large total volume
   - Extends beyond the initial surge
3. Velocity: Speed of Spread
   - Rapid acceleration to peak
   - Sharp upward trajectory
   - Quick release-to-breakout time



Key Concepts in Virality

Virality in content is the "rapid" spread of information across social networks

- Reach: The total number of people exposed to the content
- Virality Coefficient (K-factor): Measures how many new users each existing user brings. A K-factor greater than 1 indicates viral growth

$$R(t) = R_0 \times K^t$$

  $R(t)$: the reach at time t, R_0 : the initial reach, and K : virality coefficient
- Tipping Points: The critical point where content starts to spread exponentially
  - Critical mass: Minimum shares needed for exponential growth
  - Threshold models: Once enough users share content, others follow suit
- Network Effects: Content value increases with more users engaging, driving further sharing
- Positive feedback or Viral Loop: More engagement leads to more visibility
- Time to peak: The time it takes for content to reach its maximum exposure

Structural Virality

A cascade is the process by which a piece of content spreads from one user to others through shares, retweets, or reposts

A diffusion tree models this spread as a rooted tree – it captures both the structure and sequence of content spread

- The root node represents the original source or poster
- Each edge represents a direct reshare or influence from one user to another
- Nodes are users who have shared the content

Metrics for Measuring Virality

- Cascade Size: Number of users reached
  ▷ Does not capture how the content spread structurally
- Cascade Depth: Longest chain length from the root to any node
  ▷ Sensitive to outliers and ignores the rest of the tree
- Average Depth: Mean distance from root to nodes
  ▷ Fails to capture lateral spread and branching complexity
- Branching Factor: Average number of children per node
  ▷ Ignores how branching varies and overall topology

Structural Virality Metric

Given a diffusion tree T on n nodes, structural virality $V(T)$ is defined as:

$$V = \frac{1}{\binom{n}{2}} \sum_{i<j} d(i,j)$$

$d(i,j)$ = shortest path length between nodes i and j in the diffusion tree

$V(T)$ measures the average distance between all pairs of nodes, reflecting the shape and viral spread of the cascade-

This metric successfully distinguishes broadcast cascades (shallow, star-like) from viral cascades (deep, multi-step) by quantifying their structural differences

Figure 1 **A Schematic Depiction of Broadcast vs. Viral Diffusion, Where Nodes Represent Individual Adoptions and Edges Indicate Who Adopted from Whom**

source: Cornell Univeristy

Diversity of Structural Virality

Structural virality varies with the size of cascades across different online platforms

Every video, news story, image, or petition on Twitter over 12 months (1.4b events)

▷ "Popular" cascades (more than 100 retweets; $\sim 350,000$ events)

Structural virality of events range from ≈ 2 ("broadcast") to $\approx \log(N)$ ("viral")

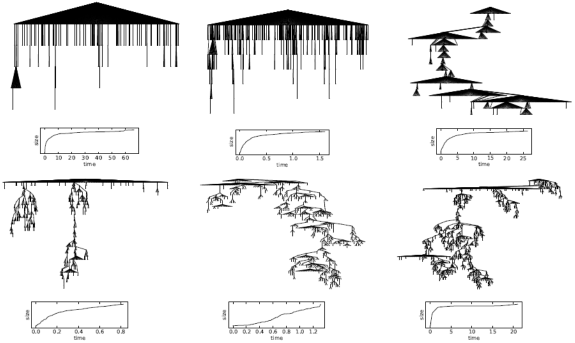
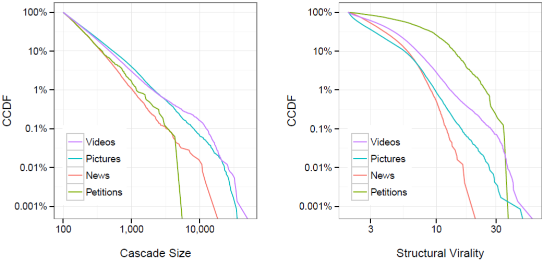


Figure 3 A random sample of cascades stratified and ordered by increasing structural virality, ranging from 2 to 50. For ease of visualization, cascades were restricted to having between 100 and 1000 adopters. Cumulative adoption curves (i.e., total cascade size over time) are shown below each cascade, with time indicated in hours. source: Goel et al. (2016) The Structural Virality of Online Diffusion

Popularity and Structural Virality

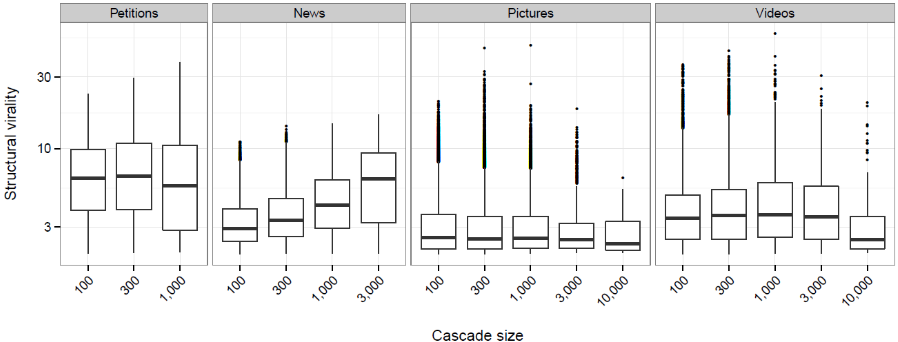


Size and structural virality distributions on a log-log scale for cascades containing at least 100 adopters, separated by domain.

source: Goel et al. (2016) The Structural Virality of Online Diffusion

- Larger cascades tend to have higher average structural virality, but popularity alone is a poor proxy for virality
- Many large cascades remain broadcast-like with shallow, star-shaped diffusion trees
- Some smaller cascades exhibit relatively high structural virality, indicating deep multi-step diffusion despite limited size
- Demonstrates that content can be popular without being truly viral in terms of diffusion structure
- Size and structural virality are related but distinct dimensions

Popularity and Structural Virality



Boxplot of structural virality by size on a log-log scale, separated by domain. Lines inside the boxes indicate median structural virality, while the boxes themselves show interquartile ranges. source: Goel et al. (2016)

Core Drivers of Virality and Viral Marketing

Distinction Between Viralization and Traditional Marketing

Viralization differs from traditional marketing in several ways:

- User-driven: Viralization relies heavily on organic sharing and user interactions, whereas traditional marketing is often push-driven (e.g., advertisements)
- Exponential growth: Viral content spreads exponentially, while traditional marketing typically sees linear growth based on campaign spending
- Low cost: Viral campaigns can spread with minimal cost, whereas traditional marketing often requires significant budget allocations
- Unpredictability: Viralization can be difficult to predict and control, while traditional marketing efforts can be carefully planned and executed

Viral Marketing and Social Media Advertising

In digital marketing, virality refers to content's ability to spread rapidly across networks. Key drivers of virality in social media advertising:

- Emotional triggering Content: Simple, humorous, joyful, empathy arousing, and urgent messaging to encourage interaction
- Influencer Marketing: Value of posts grows as more people interact and share them
- Relevance and timeliness of content
- High-quality, visually appealing media
- Monitor Brand Persona: Understand how the public perception of a brand, including public figures, affects marketing outcomes
- User Data Tracking: Analyzing engagement metrics (likes, comments, shares) helps improve future ad performance
- Diversify Marketing Channels: Combine social media with other marketing strategies (both online and offline)

Creating Viral Content and Maximizing Virality

Successful viral content includes:

- Surprise and unexpected twists to intrigue the audience
- High emotional resonance, with content evoking joy, humor, or awe
- Storytelling to build connection and engagement with the audience
- Humor and empathy, creating universal appeal and relatability
- Clear visuals and easy-to-digest formats, like lists or infographics

Effective viral marketing strategies:

- Hook, Line, Sinker Strategy:
  - Hook: A catchy headline that grabs attention
  - Line: Engaging content that delivers the promise of the headline
  - Sinker: Make the content shareable and memorable
- Influencer collaborations: Leverage influencers to amplify reach
- Platform-specific content: Tailor content to fit the algorithm and user behavior of each platform (e.g., TikTok, YouTube, Instagram)

Viral Campaign Examples and Case Studies

Notable viral marketing successes:

- MrBeast: Leveraged extreme challenges and stunts to drive views, using the "shock value" and unique content
- TikTok Challenges: Short-form, engaging videos that tap into trends and evoke emotion
- Old Spice Campaign: Mixed humor, surprise, and storytelling with the "The Man Your Man Could Smell Like" series

These examples illustrate how creativity, emotional appeal, and trend awareness drive virality

Risks and Limitations of Social Media Advertising

Despite the potential for success, viral marketing comes with challenges:

- Negative Information Spread: The rapid spread of both positive and negative information can impact brand reputation

  Example - Tesla: Elon Musk's tweet about selling Tesla stock caused a significant drop in stock prices, showing how easily social media can influence public perception

- Unpredictability: Virality is not guaranteed, even with perfect strategies
- Reputation risks: Content can quickly spiral out of control, especially if controversial
- Saturation: Over-reliance on viral tactics can exhaust audiences and reduce engagement
- Over-reliance on Social Media: Companies that focus too heavily on social media for marketing, like Groupon, have faced instability

Drivers of Viral Phenomena

The success of viral content depends on a combination of:

1. Content Appeal: Engaging and emotionally resonant content triggers sharing
2. Network Effect: Social networks amplify content spread through user interactions (share, repost, like)
   - Direct sharing: Social platforms allow users to directly share content with their friends and followers, instantly amplifying its reach.
   - Interconnected communities: Viral content spreads faster in tightly knit communities
   - Cascade effects: As one user shares content with their network, those users further share it with their own networks, creating a viral cascade
   - Influencer impact: Influencers with large followings can dramatically increase the speed and scale of viral spread through a single share
3. Algorithmic Amplification: Platforms like YouTube and TikTok boost content through recommendation algorithms based on user engagement
   - Continuous engagement: interaction $\iff$ promotion
   - Positive feedback loop:
   - Personalized recommendations:

Key Characteristics of Viral Content

The key characteristics of viral content include:

- Emotional appeal: Evokes strong emotions, such as joy, surprise, or anger
- Relatability: Resonates with a broad audience or specific communities
- Brevity and clarity: Concise and easily consumable- suitable for rapid sharing
- Shareability: Easily shareable, either through platform design (e.g., share buttons) or due to its intrinsic qualities (e.g., humor), spreads quickly
- Timeliness: Often capitalizes on current trends or timely events

Role of Production Quality and Creativity in Viral Success

- Visual appeal: High-quality, clear resolution, good lighting, and editing
- Creativity: Unique and creative concepts standing out from the crowd
- Polished and professional content: Credible and engaging $\Longrightarrow$ high ranking
- Storytelling: Well-crafted narratives or emotional appeals

YouTube's Recommendation System

YouTube Evolution and growth

- 2005: Founded and launched video sharing globally
- 2006: Acquired by Google, enabling rapid scaling
- 2007-2010: Partner Program introduced, monetizing content
- 2010-Present: Expanded with live streaming, YouTube Shorts, and premium services

Why YouTube stands out:

- Video-Centric: Focuses exclusively on video, offering optimal tools for production and distribution
- Monetization: Structured system with ads, memberships, and live features (compared to TikTok, Instagram)
- Long-Form Content: YouTube excels with long-form videos
- Search Integration: Google search integration enhances video discoverability

Comprehensive video ecosystem and monetization offer an edge over others

YouTube's Market Dominance & Growth

- Scale: 2+ billion monthly active users, second most visited site after Google
- Diverse Content: Educational to entertainment cater to wide audience
- Monetization: Ads, memberships, and live streams offer creators revenue
- Algorithm: Recommender increase engagement and watch time
- Video embedding: Embedding videos on websites and blogs, expands content reach beyond the platform
- Community building: Comment and rating systems fostered community engagement and active participation
- Growth through virality: Viral videos e.g., "Evolution of Dance" and "Charlie bit my finger" helped rapid growth as they spread via social networks and email
- Cross-platform integration: YouTube integrates seamlessly with other Google services, including Google Ads, enhancing its reach
- Scalability: The platform handles vast amounts of content while maintaining high-speed video streaming across the globe

YouTube Recommendation System Overview

YouTube's recommender is central to its viralization mechanisms. It analyzes user behavior to suggest videos, to increase user engagement. Key features include:

- Personalized Suggestions: Recommends content based on user behavior and interests
- Top N Recommendations: Provides a ranked list of top recommended videos based on user preferences
- Personalized Thumbnails: Eye-catching thumbnails increase click-through rates (CTR).
- Trending Section: Popular videos based on views, likes, and engagement are highlighted
- Algorithmic Feedback Loop: Highly engaging videos get more visibility, leading to further engagement.

Natural Viral Loop: Popular $\Longrightarrow$ More recommended $\Longrightarrow$ More views $\Longrightarrow$ More popular

Key Components of YouTube's Algorithm

YouTube's algorithm uses multiple factors to rank and recommend videos:

- Click-Through Rate (CTR): Thumbnails and titles that encourage clicks improve ranking
- Engagement Signals: Likes, shares, comments to gauge interest
  - ▷ Social proof: Video with more engagement is perceived as valuable
- Views: Immediate and publicly accessible indicator, influence recommendation
- Watch Time: Time users spend watching a video, a strong quality signal
- Retention: Videos keeping viewers engaged for longer periods rank higher
- Session duration: Users continue watching other recommended videos
- Content Relevance: Metadata like titles, descriptions, and tags help align videos with user queries
- Video Popularity: Trending and highly engaging videos get promoted
- Diversity: Recommends mixed content, balancing familiar and new content
- Co-visitation Count: Tracks user behaviors to recommend related videos
- Related Videos Graph: Models relationships between videos

Co-Visitation Count

YouTube's recommender leverages co-visitation counts and related videos graph:

Co-Visitation Matrix: Tracks how often two videos are watched together. High co-visitation increases the likelihood of their recommendation

co-visitation count between videos A and B $$C(A,B) = \sum_{i=1}^{N} f_i(A,B)$$

▷ $f_i(A,B)$: 1 if both videos are watched in session i, otherwise 0, N : number of user sessions

- Temporal correlation: Takes time proximity into account, ensuring relevance for short-term trends
- Recommendation relevance: The higher the co-visitation count, the more likely video B will be recommended after video A

Session-Based Recommendations: Predicts the next likely video based on real-time user behavior and session patterns

- Session continuation: Recommending co-visited videos ensures viewers stay on the platform longer, leading to increased watch time

Example: After watching a music video, YouTube may suggest another video often watched in the same session by users with similar preferences

Related Videos Graph

The Related Videos Graph helps recommend videos based on their similarity

- Nodes represent individual videos
- Edges represent relationships between videos based on co-visitation counts or user interactions
- Weight of edges is determined by the strength of the relationship, often derived from co-visitation counts or engagement metrics

This graph helps YouTube generate recommendations by finding and ranking neighboring nodes (videos) for a given video
Ranking videos in YouTube's Related Videos Graph uses algorithms like:

- PageRank: Scores nodes based on number and quality of co-visitation links
- Personalized PageRank: Adjusted PageRank based on user-specific preferences and watch history
- Random Walk with Restart: Explores the graph dynamically, restarting periodically to prioritize related content

Top N Recommendations & Ranking

YouTube generates Top N recommendations that are dynamically refreshed based on changes in user behavior:

- Candidate Generation: A broad selection of potential recommendations is generated
- Relevance Scoring: Candidates are ranked based on factors like co-visitation, engagement, and watch history
- Top N Selection: The highest-scoring videos are shown to the user in descending order

Methods for Generating Top N Recommendations

Top N recommendations are generated using various methods, including:

- Collaborative filtering: Suggests videos based on the behavior of similar users (e.g., users who watched video A also watched video B)
- Content-based filtering: Recommends videos that are similar in content or metadata (e.g., genre, topic) to videos the user has watched
- Hybrid models: Combines collaborative and content-based filtering for more accurate recommendations

Example: YouTube may recommend similar tutorial videos based on both the content of previously watched videos and the preferences of other users

Algorithmic Optimization for Personalized Recommendations

- Real-time feedback: The algorithm adapts in real time based on user interactions (likes, dislikes, watch time) to refine future recommendations
- Exploration vs. exploitation: Balances recommending familiar content (exploitation) with new videos (exploration) to prevent user fatigue
- User Embeddings: Deep learning models create a numerical representation (embedding) of each user's preferences to match them with relevant content

Personalization and Deep Learning

YouTube integrates deep learning for enhanced personalization:

- Deep Neural Networks (DNN): Used for video ranking and personalized recommendations based on complex user behavior
- Reinforcement Learning (RL): Continuously adjusts recommendations based on real-time user feedback, improving accuracy and relevance
- Dynamic Adjustments: The recommendation system dynamically adjusts in response to evolving user preferences and content trends

Cold-Start Problem & Hybrid Approach

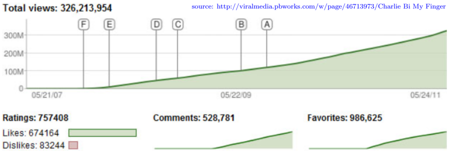
YouTube addresses the cold-start problem with hybrid techniques:

- Content-Based Filtering for New Videos: Recommends videos based on metadata when there is little user interaction
- Cross-Platform Data: Leverages data from other Google services to boost recommendations for new users or videos
- Hybrid Approach: Combines collaborative filtering and content-based techniques to mitigate the cold-start problem

Analysis of One of YouTube's Most Viral Videos

“Charlie bit my finger”

“Charlie bit my finger”, one of the most iconic viral videos on YouTube




NFT of soon-to-be-deleted ‘Charlie bit my finger’ YouTube video sells for over $760,000

By Eric Levenson, CNN

2 minute read · Updated 11:50 AM EDT, Sun May 23, 2021

- May 22, 2007: Uploaded by Howard Davies-Carr to share with family
- Feb 2008: Surpassed 2.6 million views
- Mar 2008: Reached 12 million views
- Dec 2008: 12th most-viewed video on YouTube with 65 million views
- Oct 2009: Most-viewed video on YouTube, surpassing "Evolution of Dance"
- 2011: Surpassed 500 million views, first user-generated video to achieve this
- May 2021: Sold as an NFT for over $700,000 and removed from YouTube
- Jan 2024: Returned to YouTube after being private for 30 months
- 2025: Amassed nearly 900 million views

Factors Contributing to Virality

These factors helped the video accumulate millions of views rapidly, setting a precedent for viral content on YouTube

- Short duration: The video is short, under a minute, making it easy to watch and share
- Emotional Resonance: The genuine reactions of the children evoke joy, amusement, laughter and empathy, enhancing shareability
- Humor and simplicity: The video's content is simple yet humorous, featuring a playful interaction between two young brothers
- Authenticity and Relatability: The straightforward and unscripted nature of the video made it relatable to a broad audience
- Cultural relevance: The video became a reference point for family life and relatable moments, making it resonate with a wide audience
- Media Coverage: Extensive coverage by news outlets amplified its reach

Metrics and Viewer Engagement Trends

The viral success of "Charlie bit my finger" can be analyzed through metrics:

- Views: The video garnered over 800 million views, making it one of the most-watched videos in YouTube's early history
- User engagement: accumulated millions of likes and thousands of comments, increasing its reach
- Social media sharing: High share rates on Facebook and Twitter (rising during the same period), and via email led to exponential growth in views
- Watch time: Despite its short duration, maintained high watch completion rates, signaling strong user engagement
- Virality timeline: The video experienced an initial spike in views followed by sustained engagement over several years, indicating long-term virality
- Early Adoption and Sharing: Initial sharing among early YouTube users facilitated rapid spread
- Algorithmic promotion: YouTube's recommendation algorithm favored the video due to its high engagement, further amplifying its visibility

Cultural Impact and Legacy

- Meme Culture: Inspired numerous parodies and remixes, embedding it in internet culture
- Monetization Strategies: Demonstrated the potential for creators to earn revenue through viral content
- Influence on Future Content Creation: Set a precedent for the viral potential of everyday moments

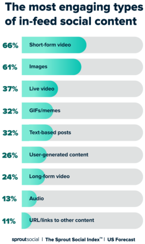
Instagram Algorithms

Signals Used by the Instagram Algorithm

The Instagram algorithm determines what content users see across different sections: Feed, Stories, Explore, and Reels. It uses various signals to predict what content will engage users and show them what they are most likely to enjoy

The goal is to enhance user experience by prioritizing relevant content and keeping users engaged. There isn't a single universal algorithm; each section has its own unique algorithm with tailored signals and rankings

- User Engagement: Likes, comments, shares, direct interactions
- Interaction History: How often you interact with the account or type of content
- Content Characteristics: Popularity, content type (video/photo), length, quality
- Timeliness: How recent the post is.
- User Activity: Posts type user interact with and frequency



Instagram Reels Algorithm

Reels prioritize entertaining, funny, and engaging videos. The Reels algorithm uses the following signals:

- User Engagement: Interactions like likes, comments, and shares with Reels
- Interaction History: Past interactions with the creator of the Reel
- Content Features: Audio or music used in the Reel, and the quality of the content
- Creator Popularity: How well-known or popular the person who posted it is

What to Avoid:

- Low-resolution videos
- Watermarked videos (e.g., TikTok logos)
- Reels without audio
- Content with too much text or political topics

Feed and Story Algorithm

Instagram Feed and Stories prioritize content from friends and people users interact with the most. Key signals include:

- Post Popularity: Number of likes and comments, post type (video/photo), and post recency
- Interaction History: How often a user has interacted with the content creator
- User Activity: Posts that reflect user interests based on previous engagements
- Content Relevance: Content that aligns with the user's interaction history

What to Avoid:

- Posting too frequently (overloading users' feeds)
- Violating community guidelines
- Posting misinformation

Engagement is a critical signal for ranking posts in both Feed and Stories.

Explore Page Algorithm

The Explore page shows content based on past interactions and what other users with similar interests are engaging with. Key signals include:

- Post Popularity: The speed at which a post garners likes, shares, and comments
- Interaction History: Previous interactions with users featured on Explore
- User Activity: Patterns of engagement, such as liking or saving posts
- Content and Account Popularity: How well posts from the same account have performed recently

What to Avoid:

- Content that violates community guidelines.
- Sensitive or inappropriate content that could be flagged.

The Explore algorithm curates content based on users' likes and interactions with similar content across Instagram.

Instagram Algorithms Updates (2022)

- Return of Photos: Instagram shifted focus back to photos, balancing with video content
- Increased In-Feed Suggestions: Suggested posts from accounts users don't follow based on similar interests
- Feed Options: Users can choose between:
  - Chronological Feed – Posts ordered by recency
  - Algorithmic Feed – Posts ordered based on the platform's algorithm
  - Favorites Feed – Content from a curated list of favorite accounts

To make the most of the Instagram algorithm, apply these strategies:

- Post Consistently: Regular posts increase visibility and engagement
- Focus on Video Content: Reels have the highest reach and engagement
- Optimize Captions and Hashtags: Use keywords and hashtags to make your content discoverable
- Encourage Engagement: Ask questions, use Story stickers, and engage your audience with interactive content
- Post at Optimal Times: Leverage analytics to post when your audience is most active

Engagement Tips for Success

Engagement (through comments, likes, and other forms of interaction) plays a significant role in how the algorithm ranks your content. Here are some ways to drive engagement:

- Spark Conversations: Use captions that encourage comments and interactions
- Use Story Stickers: Interactive elements like polls, questions, and quizzes boost engagement
- Post Polls or Trivia: Engaging posts that prompt responses
- Use "Tag a Friend" Prompts: Encourage users to tag friends, further boosting visibility

TikTok Engagement

TikTok and Its Addictive AI Algorithm

TikTok's For You Page (FYP) is driven by an AI-powered recommendation system that unlike Instagram and Facebook, doesn't prioritize follower count or social network proximity

TikTok's AI algorithm, central to its rapid rise, is highly addictive and engaging

- User-centric algorithm to personalize content and retain user attention
- TikTok emphasizes video virality, not based on followers or views, but on engaging content
  - ▷ Content discovery decoupled from network size, relying on engagement signals (likes, shares, rewatches), allowing new creators to gain viral success
- TikTok's algorithm allows content from new users to quickly go viral, regardless of follower count

Jennifer Lopez's TikTok video received 71 million views with 5 million followers, illustrating the power of TikTok's recommendation engine

TikTok's AI Recommendation System

TikTok's recommendation system has three main building blocks that interact to create a highly personalized content experience that keeps users hooked:

- Self-training AI Engine: Continuously learns from user behavior
- Content Tags: Tags content to improv discoverability and engagement
- User Profiles and Scenarios: Recommendations based on user preferences

1. Content Audit: AI-driven audit of videos to detect inappropriate or malicious content ▷ NLP and computer vision to assess video relevance and quality
2. Initial Batch Processing: Content receives initial exposure in a cold-start pool
3. Metric-Weighted Content Performance: User interactions (likes, shares, rewatches) help assign performance scores to content
4. Feedback-Based Amplification: High-performing content is further promoted to targeted user groups
5. Tailored Trending Pools: Top-tier content is promoted indiscriminately, often resulting in viral moments
6. Delayed Ignition: Older content can gain viral traction over time as it resurfaces on the FYP

Network Effect

Two types of network effect

What you do or use may depend on what other do or use

1. Information Network Effect: The fact that other do something gives you information about some thing This could lead to Information Cascade or Fallacy of Crowd
2. Valuation Network Effect: The fact that other use certain product makes the product valuable e.g., The only person having a fax machine will have no value for it, it only have value if others also use it

1. Population Based: Number/fraction of people adapting a product or service
   - Deterministic Interaction
     - Information Cascade and Synchronization
   - Random Interaction
     - Tipping Model or Threshold models: Users have an exposure threshold they share. Once a critical mass is reached, the content spreads rapidly
     - Diffusion models: Simulate how content spreads from person to person across a network. In SI model (Susceptible-Infected) users become "infected" (i.e., share the content) once exposed
2. Topology Based: Underlying graph topology/connectivity plays a role

Population Based and Topology Based Network Effect Models

Table 7.1 Influence Models in Chapters 7 and 8. Tipping and contagion models are highly related, as are information cascade and random walk, and diffusion and infection. Those models we discuss in Long Answers are in italics, while the other models are in Advanced Material. The precise forms of deterministic or random interaction models will be presented as we go through these two chapters.

	Deterministic Interaction Model	Random Interaction Model
Population based	*Information cascade* Synchronization	*Tipping* Diffusion
Topology dependent	*Contagion* Random walk	*Infection*

Which model to use, depends on how people respond in the system, if all people respond at the same time you may use synchronization model, if a few early adapters change others minds, you may use diffusion or contagion models if a few people carry the system over some threshold, then we may use tipping models.

Sequential Decision & Information Cascade Model

Information Cascade Model

The Information Cascade Model explains how individuals make decisions based on the actions of others, leading to a group behavior that might not align with private information

Individuals decide whether to share a piece of news based on how many previous users have shared it:

- Each user sees the number of prior shares
- Users decide to share if they see more than 5 previous shares, mimicking the cascade effect
- This decision-making process can lead to widespread sharing of news, regardless of its accuracy

- Overview of the model's relevance in understanding social media trends and economic decisions
- Theoretical foundation: sequential decision-making under uncertainty
- Impact of observing others' decisions on personal choice

Salganik, Dodds, and Watts (2005) Experiment

SDW studies the role of social influence on individual decision-making (preference shaping) and how it leads to collective behavior (e.g., viral content)

Experiment Study of influence of others' actions on users songs preferences

- 14,341 participants rated/downloaded 48 songs from unknown bands
  - Participants rated each song (1 to 5 stars) while listening to it
  - After rating, they were asked if they would like to download the song
- Four different variations based on two factors:
  1. Order of Song Presentation: Songs were presented either at random or in descending order of current download counts
  2. Visibility of Download Numbers: The current download numbers for each song were either shown or hidden
- Participants song ratings and choices were noted

Demonstrated that individual choices could be strongly influenced by the choices of others, even if those choices were made without full information

Experiment 1 and Experiment 2 Setup

Experiment 1:

- Songs displayed in a 3-column grid, order randomized for each participant
- In the social influence condition, participants could see the download counts for each song as they chose
- In the independent condition, participants saw the same song grid, but no download counts were shown

Experiment 2:

- Songs displayed in a 1-column list, ordered by descending download counts
- In the social influence condition, the download counts were visible
  - ▷ reinforcing social influence by showing which songs were most popular
- In the independent condition, the songs were displayed in random order, and download counts were hidden

Experiment	Participants	Independent condition	Per 8 social influence world
Experiment 1	7,149	1,400	700
Experiment 2	7,192	1,446	700

In Social influence condition, participants were randomly assigned to one of 8 different "worlds"

Market Share and Gini Coefficient

The market share m_i of song i in an experiment (world)

$$m_i = \frac{d_i}{\sum_{k=1}^{S} d_k}$$

- d_i : number of downloads for song i
- S : total number of songs in experiment

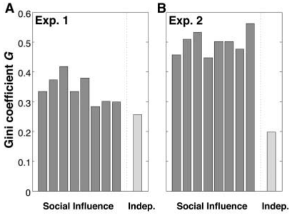
The Gini coefficient G measures the inequality of success among songs

$$G = \frac{1}{S^2} \sum_{i=1}^{S} \sum_{j=1}^{S} |m_i - m_j|$$

- G ranges from 0 (complete equality)
- to $G = 1$ (maximum inequality)

The experiment showed that social influence increases the Gini coefficient, meaning more popular songs became even more popular

- G increases with social influence
- Songs that are already popular gain even more popularity
- less popular songs stay ignored



Unpredictability Measurement

Unpredictability is measured by variation in a song's market share across worlds

The unpredictability u_i for song i is :

$$u_i = \frac{1}{\binom{W}{2}} \sum_{j=1}^{W} \sum_{k=j+1}^{W} |m_{i,j} - m_{i,k}|$$

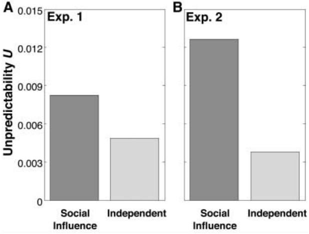
- W is the number of worlds
- $m_{i,j}$ and $m_{i,k}$ are the market shares of song i in worlds j and k

Overall unpredictability U is the average unpredictability across all songs:

$$U = \frac{1}{S} \sum_{i=1}^{S} u_i$$

The experiment showed that Social influence increases randomness in the success outcomes of songs

- U increased with social influence
- especially in Experiment 2 when the social signal was stronger
- less popular songs stay ignored



Findings and Implications

Social influence causes a shift in the perceived popularity of songs

- Social influence significantly alters individual decision-making, leading to group behavior even when participants did not know why they favored certain songs
- It demonstrated that popularity is not purely based on intrinsic quality, but can be heavily influenced by external social factors
- The experiment revealed the role of unpredictability, where songs in the middle range of ratings were much more susceptible to the influence of others

This experiment helps understand content virality on platforms e.g., YouTube

If a video gains initial views, it is more likely to be recommended to other users, thus increasing its viewership

▷ the "rich get richer" effect

This creates a feedback loop, similar to the social influence observed in the experiment, where the popularity of a video continues to grow due to the influence of others

Following the Crowd

Social relations between people influence each other's behavior and decisions

- opinions, political positions they hold
- activities, technologies they adopt

- Say you have done your own research and chosen restaurant A for dining
- Arriving there, you see restaurant A is empty and restaurant B is full
- It may be rational to join the crowd at B rather than follow your research

  ▷ If each diner has independently decided that B is better, then collectively this may be stronger than your own research

In this case, *herding* or an *information cascade* has occurred

▷ behaviors that cascade from one node to another like an epidemic! and produce collective outcomes – We give rational reasoning for this behavior

Sequential Decision Making

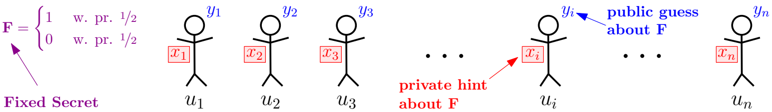
- Individuals make choices sequentially
- Private Signals: Individuals receive unique private information/signals
- Public Actions: They observe others' actions and adjust their behavior
- Each individual observes all previous public actions but not the private signals that informed those actions

▷ From others public actions people infer what they know

- This creates an information dependence and leads to an information cascade – individuals, influenced by the actions of others, ignore their private information and make decisions based on observed behavior
  - Not mindless imitation or social pressure to conform! It is making rational inferences from limited information
- Inference makes information cascade, benefit to individuals makes direct-benefit cascade

▷ e.g., fax machines. A fax machine is only useful and beneficial if your collaborators possess a fax machine. No information about fax machines is needed.

Sequential Decision Making: A thought Experiment



- Consider a scenario where individuals are sequentially making decisions about a fixed, secretly chosen number $F \in \{0, 1\}$ (e.g., a binary event)
- Each user u_i gets a private signal, $x_i \in \{0, 1\}$ (a probabilistic version of **F**)

$$Pr[\ u_i\text{'s private signal is correct}] \ = \ Pr[x_i = \mathbf{F}] \ = \ p_i \ > \ 1/2 \quad \text{(say 60\%)}$$

- u_i uses x_i and $y_1, y_2, \ldots, y_{i-1}$ to announce his guess $y_i \in \{0, 1\}$ for **F**
  ▷ Sequential public announcement of his decision
- u_i rationally decides y_i ▷ maximizing probability of correct guess
- Assume $p_1, p_2, \ldots, p_n$ are all equal to $p > 1/2$

Goal: Explore how users adjust their decisions based on decisions of others

SDM: First Person

What should the first person do? u_1, has only their private signal x_1

If $x_1 = 1$, announce $y_1 = 1$ If $x_1 = 0$, announce $y_1 = 0$

$Pr[x_1 = 1] = Pr[x_1 = 1|\mathbf{F} = 1]Pr[\mathbf{F} = 1] + Pr[x_1 = 1|\mathbf{F} = 0]Pr[\mathbf{F} = 0] = (p)\frac{1}{2} + (1-p)\frac{1}{2} = \frac{1}{2}$

$$Pr[\mathbf{F} = 1|x_1 = 1] = \frac{Pr[\mathbf{F} = 1] \times Pr[x_1 = 1|\mathbf{F} = 1]}{Pr[x_1 = 1]} = \frac{\frac{1}{2} \times (p)}{\frac{1}{2}} = p > \frac{1}{2}$$

Since $Pr[\mathbf{F} = 1|x_1 = 1] > 0.5$ (60%), it is rational to announce $y_1 = 1$

- Similarly, $Pr[\mathbf{F} = 0|x_1 = 0] = p > 1/2$
- $Pr[\mathbf{F} = 1|x_1 = 0] = 1 - p < 1/2$
- $Pr[\mathbf{F} = 0|x_1 = 1] = 1 - p < 1/2$

The first person's decision is purely based on their private signal, as no one else has made a decision yet

SDM: Second Person

What should the u_2 do? u_2 has their private signal x_2 and also observe y_1

u_2 cannot see x_1 but can infer that x_1 must be equal to y_1

- if both x_1 and x_2 are 0, then $\mathbf{F}$ is more likely to be 0, so announce $y_2 = 0$
- Similarly, if both x_1 and x_2 are 1, then y_2 should be 1
- If one is 0 and the other is 1, then randomly choose y_2 to be 0 or 1

$$Pr[x_1, x_2 = 0, 0] = Pr[x_1, x_2 = 0, 0|\mathbf{F} = 1]Pr[\mathbf{F} = 1] + Pr[x_1, x_2 = 0, 0|\mathbf{F} = 0]Pr[\mathbf{F} = 0] = (1-p)^2\tfrac{1}{2} + p^2\tfrac{1}{2}$$

$$Pr[\mathbf{F} = 0|x_1, x_2 = 0, 0] = \frac{Pr[\mathbf{F} = 0]Pr[x_1, x_2 = 0, 0|\mathbf{F} = 0]}{Pr[x_1, x_2 = 0, 0]} = \frac{Pr[\mathbf{F} = 0]Pr[x_1 = 0|\mathbf{F} = 0]Pr[x_2 = 0|\mathbf{F} = 0]}{Pr[x_1, x_2 = 0, 0]}$$

$$Pr[\mathbf{F} = 0|x_1, x_2 = 0, 0] = \frac{\frac{1}{2}p^2}{\frac{1}{2}(1+2p^2-2p)} > \frac{1}{2} \quad \text{when } p > 1/2$$

- Similarly, $Pr[\mathbf{F} = 1|x_1, x_2 = 1, 1] > 1/2$
- $Pr[\mathbf{F} = 1|x_1, x_2 = 0, 0] < 1/2$
- $Pr[\mathbf{F} = 0|x_1, x_2 = 1, 1] < 1/2$

SDM: Third Person

u_3 has their own private signal x_3 and also observes two public actions y_1 and y_2

- If $y_1 = y_2 = x_3$, then u_3 will obviously pick that number
- If $y_1 = y_2$ but x_3 differs, you still will pick what 2 public signals suggest
- If $y_1 \neq y_2$, then public actions by prior users collectively do not tell you anything, and you should just rely on your own private signal
  - u_3 is in same situation as u_1. u_4 will be in same situation as u_2

Denote $y_1 = y_2 = 1$ and $x_3 = 0$ as $(1, 1, 0)$

What is the probability that $\mathbf{F} = 1$ given this sequence of $(1, 1, 0)$?

$(y_1, y_2, x_3) = (1, 1, 0)$ results from $(x_1, x_2, x_3) = (1, 1, 0)$ or $(1, 0, 0)$ but u_2 randomly chose $y_2 = 1$

$$P[\mathbf{F} = 1|(1,1,0)] = \frac{P[\mathbf{F} = 1]P[(1,1,0)|\mathbf{F} = 1]}{P[(1,1,0)]} = \frac{P[\mathbf{F} = 1]P[(1,1,0)|\mathbf{F} = 1]}{P[\mathbf{F} = 1]P[(1,1,0)|\mathbf{F} = 1] + P[\mathbf{F} = 0]P[(1,1,0)|\mathbf{F} = 0]}$$

$$P[\mathbf{F} = 1|(1,1,0)] = \frac{0.5\left(p^2(1-p) + 0.5p(1-p)^2\right)}{0.5(p^2(1-p) + 0.5p(1-p)^2) + 0.5((1-p)^2p + 0.5p^2(1-p))} = \frac{2p + (1-p)}{2p + 1 - p + 2(1-p) + p} = \frac{1+p}{3} > \frac{1}{2}$$

By this point, the individuals have started to disregard their own signals and follow the public action, leading to an information cascade

SDM: Information Cascade

If first two guesses are the same, the third should make the same guess as the first two, regardless of his own private info! An information cascade has begun!

The Fourth Participant and Onward in the cascade case!

- First two guesses were the same, say $(y_1, y_2) = (1, 1)$

  ▷ 3^{rd} guess has to be 1 too

- 4^{th} participant, observed $(y_1, y_2, y_3) = (1, 1, 1)$
  - First 2 guesses conveyed perfect info
  - 3^{rd} guess conveys no info– It has to be 1, no matter what u_3 gets
- 4^{th} is in same situation as 3^{rd}! – should guess 1 regardless of what u_4 sees
- This will continue– If first 2 guesses were 1, then all subsequent participants will guess 1!

Once an odd-numbered user and the next even-numbered user show the same public action, the next user will just follow, no matter what her private signal is. An information cascade thus starts – The cascade continues as each subsequent person follows the majority's choice, disregarding their private information

Probability of Cascade (Up or Down)

A cascade occur depending on the private signals and the previous decisions made in the sequence – A cascade can go in one of two directions:

- UP-CASCADE: The group decides that $F = 1$ – all subsequent users choose 1
- DOWN-CASCADE: The group decides $F = 0$ – all subsequent users choose 0

No cascade occurs after two public actions if the first two decisions are different

Pr[No Cascade after 2 people made their public actions] $=$ Pr[the first 2 private signals are different]

If $\mathbf{F} = 1$, then $Pr[x_1 = 1, x_2 = 0] = p(1-p)$ and $Pr[y_2 = 0] = 1/2$

$Pr[y_1 = 1, y_2 = 0] = p(1-p)/2$ and $Pr[y_1 = 0, y_2 = 1] = p(1-p)/2$

$Pr[\text{no cascade when } \mathbf{F} = 1] = Pr[y_1 \neq y_2] = p(1-p)$ and by symmetry

$Pr[\text{no cascade when } \mathbf{F} = 0] = Pr[y_1 \neq y_2] = p(1-p)$

In conclusion, $Pr[\text{no cascade}] = p(1-p) = \textsc{ProbNo}$

$$Pr[\textsc{up-cascade}] = Pr[\textsc{down-cascade}] = \frac{1 - \textsc{ProbNo}}{2}$$

Cascade Dynamics & Probabilities for Multiple Users

The probability of no cascade after $2n$ users have made decisions is:

$$P[\text{no cascade after } 2n \text{ actions}] = (\textsc{ProbNo})^n = (p(1-p))^n$$

While the probability of an up or down cascade after $2n$ actions is:

$$Pr[\textsc{up-cascade}] = Pr[\textsc{down-cascade}] = \frac{1-(p(1-p))^n}{2}$$

$\because$ no cascade occurs if each pair of $(1,2),(3,4),(5,6),\ldots$ take different actions

As n increases, the probability of a cascade occurring becomes very likely – Cascades are inevitable with a sufficiently large number of individuals

The probability of correct vs. incorrect cascades depends on the value p

If the signal is very noisy, i.e., $p = 0.5$, correct cascades are as likely to happen as incorrect cascades.

If the chances of the first two participants private info being correct is small, all others will wrongly guess!

As p approaches 1, the chance of a correct cascade goes up quickly towards 1

Cascade Dynamics & Probabilities for Multiple Users

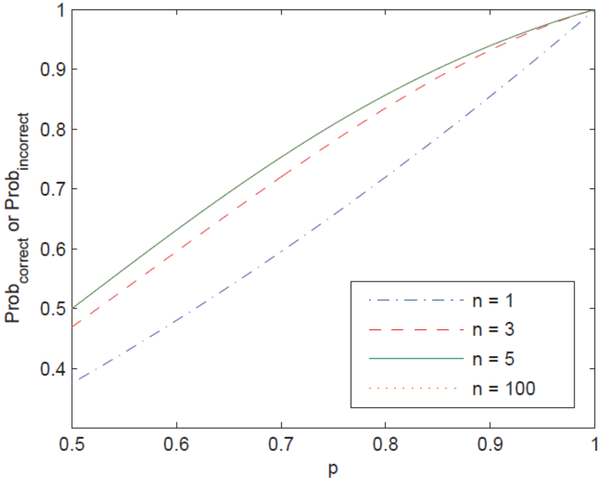


Figure 7.4 Probabilities of information cascade. As the probability p (of each person getting the correct private signal) increases, the probability of forming a correct cascade increases roughly linearly. Once the number of people n reaches 5, the curves are almost identical even as n increases further.

Breaking a Cascade

How long will a cascade last? Well, forever, unless there is some kind of intervention, e.g., a release of private signals – Some (but not all) cascades can be very fragile!

Even a little bit of that often suffices because, despite the number of people in the cascade, they all know they are just following a very small sample

- Suppose first 2 guesses are 1
- u_i and u_{i+1} get $(x_i, x_{i+1}) = (0, 0)$ and "show" it to others!
- u_{i+2} has four pieces of insider info:
  - $u_1(1)$, $u_2(1)$, $u_i(0)$, $u_{i+1}(0)$
- u_{i+2} should decide based on his/her own private signal x_{i+2}

This is the counterpart of an easy (and possibly wrong) cascade: the Emperor's New Clothes effect. Sometimes, it only took one kid's shouting out a private signal to stop the cascade

Different probabilities of Private Signals Being Correct

We assumed that everyone has the same precision/quality of private signal p

Each person i may have a different p_i

Suppose all people know the values of $\{p_i\}$ for everyone

Subsequent users may trust some preceding users more than others

Where should the high precision person be placed in the sequence of people (if you could control that)?

- Put a highest precision person earlier, a cascade may start right after her
- A higher precision person later may easily break/reverse the cascade
- Private signals keep getting overwhelmed by public actions – The impact of high precision later in the sequence reduces

Partial Observations of Previous Public Actions

What if people do not observe all the decisions made by $\{u_1, u_2, \ldots, u_{i-1}\}$ but only some of them?

u_i observe a subset of public actions $\{y_1, y_2, \ldots, y_{i-1}\}$

What if u_i only observe a summary statistics of $\{y_1, y_2, \ldots, y_{i-1}\}$

For example, number of downloads

General Cascade Model

The cascade model explains how individual decisions, made sequentially, are influenced by others' actions

Consider people $(u_1, u_2, u_3, \dots)$ sequentially making decisions to **a**ccept or **r**eject some option

- adapting a new tech
- voting for a candidate
- watching a video

The model contains the following elements:

- Initially the world is in one of two states
  - G : where the option is a good idea, $Pr[G] = p$
  - B : where the option is a bad idea, $Pr[B] = 1 - p$
- Individuals receive payoffs depending on their decision
  - If **r**eject, payoff $= 0$
  - If **a**ccept and the option is good, payoff $= v_G > 0$
  - If **a**ccept and the option is bad, payoff $= v_B < 0$
- Expected payoff with no extra information is 0 $p \times v_G + (1-p) \times v_B = 0$
- u_i gets a private signal $\in \{L, H\}$ before deciding that tends to be correct
  - H : Points to G (accepting is a good idea), $Pr[H|G] = q > 1/2$
  - L : Points to B (accepting is a bad idea), $Pr[L|G] = 1 - q$

Decision making with one signal: First Person

Individuals decide based on (1) their private signal, and (2) observation of the earlier decisions made by other individuals

u_1 only has a private signal

If u_1's private signal is H then their payoff shifts from:

$$pv_G + (1-p)v_B = 0 \qquad \text{to} \qquad v_G Pr[G|H] + v_B Pr[B|H]$$

$Pr[G|H] = \frac{Pr[G]Pr[H|G]}{Pr[H]} = \frac{p \times q}{Pr[H]}$, $\quad Pr[B|H] = \frac{(1-p)(1-q)}{Pr[H]}$, $\quad Pr[H] = Pr[H|G]Pr[G] + Pr[H|B]Pr[B]$

The new expected payoff is: $\dfrac{v_G(pq) + v_B(1-p)(1-q)}{pq + (1-p)(1-q)} > 0$

Since $q > 1/2$ and $(1-q) < 1/2$, the new payoff is greater than 0

- So the individual should accept the option when getting a high signal!

Decision making with multiple signals

Individuals decide based on (1) their private signal, and (2) earlier decisions

u_i gets a sequence S of interleaved **a**ccept and **r**eject signals **(including H/L)**

We can derive the following facts:

1. $Pr[G|S] > Pr[G]$ when $\mathbf{a} > \mathbf{b}$
   - ▷ posterior probability higher than prior probability
2. $Pr[G|S] < Pr[G]$ when $\mathbf{a} < \mathbf{b}$
   - ▷ posterior probability lower than prior probability
3. $Pr[G|S] = Pr[G]$ when $\mathbf{a} = \mathbf{b}$

Therefore the individual should

- accept when $\mathbf{a} > \mathbf{b}$
- reject when $\mathbf{a} < \mathbf{b}$
- Randomly choose when $\mathbf{a} = \mathbf{b}$

$Pr[S|G] = q^a(1-q)^b$ ▷ $\because$ signals are independent

$Pr[S] = Pr[S|G]Pr[G] + Pr[S|B]Pr[B] = q^a(1-q)^b \times p + q^b(1-q)^a \times (1-p)$

$$Pr[G|S] = \frac{Pr[G] \times Pr[S|G]}{Pr[S]} = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

Compare $Pr[S|G]$ to p using $q > 1/2$ when $a = b$ $Pr[S|G] = p$

Decision making with multiple signals

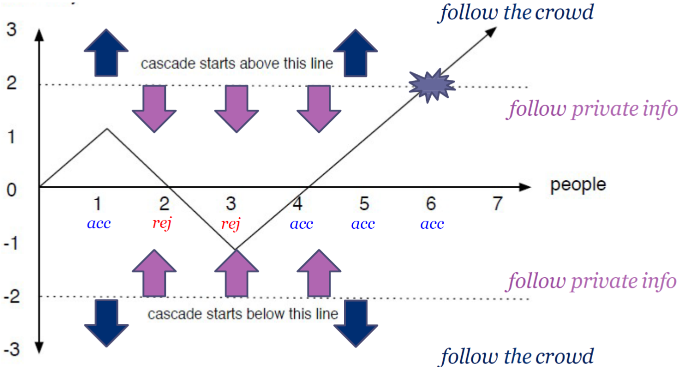
u_i gets a sequence S of interleaved **a**ccept and **r**eject signals **(including H/L)**

1. $Pr[G|S] > Pr[G]$ when $\mathbf{a} > \mathbf{b}$
   - ▷ posterior probability higher than prior probability
2. $Pr[G|S] < Pr[G]$ when $\mathbf{a} < \mathbf{b}$
   - ▷ posterior probability lower than prior probability
3. $Pr[G|S] = Pr[G]$ when $\mathbf{a} = \mathbf{b}$

Therefore the individual should

- accept when $\mathbf{a} > \mathbf{b}$
- reject when $\mathbf{a} < \mathbf{b}$
- Randomly choose when $\mathbf{a} = \mathbf{b}$

source: Hadi Amiri @ UMass-Lowell



Implications of Information Cascade

We assumed that everyone acts rationally. What each person should do can be quite different from what she actually does in real life.

There are many implications of this information cascade model, since many social phenomena exhibit the features of

- crowd following and ignoring individual's private information
- yet the chain reaction is fragile

- Social networks: rapid spread of information or misinformation
- Economics: market trends driven more by perception than fundamentals
- Politics: how public opinion can be swayed by visible majority actions
  - Partly why U.S. presidential primary elections use Super Tuesday to avoid sequential voting
  - why teenagers tend to obtain information from experiences of peers
  - why once people (correctly or incorrectly) suspect the underlying true signal has changed, the ongoing cascade can quickly reverse

TikTok and Information Cascades

TikTok and Information Cascades

TikTok's recommendation algorithm differentiates itself from traditional social media platforms by using a decentralized, sequential information cascade:

- TikTok evaluates how previous users engage with a video (view time, rewatching, etc.) and predicts whether you will enjoy it
- Content is shown to a small sample group first, and if the video is well-received, it is promoted more widely
- Unlike typical network models, TikTok's system does not depend solely on user networks; it heavily relies on engagement metrics

Through this cascade, users are shown content tailored to their preferences, rapidly creating niche clusters

Retweet Cascades on Twitter: Distribution and Structure

A retweet cascade refers to the spread of a tweet as it is retweeted by others. The cascade's size refers to the number of people reached, and the depth refers to how far the cascade falls from the original tweet.

Key findings from a year-long dataset of over 1 billion tweets:

- 93% of tweets received no retweets.
- 5%, 0.9%, and 0.3% of tweets had 1, 2, and 3 retweets, respectively
- A small percentage of tweets created large cascades, following a power law distribution

The study found that the popularity of tweets and the structure of their cascades follow this skewed distribution

Cascade Structure: Broadcast vs. Viral

Retweet cascades can be classified into two types:

- Broadcast: Everyone retweets the tweet from the same origin, leading to a cascade where the average shortest path length between retweets is about 2
- Viral: Few retweets at each step, but the tweet spreads widely. The shortest path length follows $\log(n)$, where n is the number of retweets.

The average shortest path length quantifies the structure, allowing differentiation between broadcast and viral cascades.

Figure 1 **A Schematic Depiction of Broadcast vs. Viral Diffusion, Where Nodes Represent Individual Adoptions and Edges Indicate Who Adopted from Whom**

source: Cornell Univeristy

Popularity of a tweet does not correlate with its virality; popular tweets are often associated with broadcast cascades

Application of Graph Theory and Key Insights

Graph theory plays a crucial role in analyzing retweet cascades:

- Cascade size: Number of people reached
- Cascade depth: Number of steps from the original tweet to the farthest retweet
- Shortest path lengths: To distinguish between viral and broadcast cascades

Key Insights on how information spreads on social networks:

- The distribution of retweet cascades follows a power law, with a few tweets receiving most of the engagement
- Broadcast cascades typically result in higher popularity for tweets, while viral cascades can still lead to broad reach with fewer retweets
- The structure of retweet cascades is essential for understanding the dynamics of social media virality