Social Computing and User-generated Content

ARPITA GHOSH
Cornell University

ACM EC 2013 TUTORIAL

June 16, 2013

Introduction

Contributions from users increasingly central to Web

- Information: Prediction markets, ratings, opinion polls, ...
- Content: User-generated content (UGC)
  - Reviews, Q&A forums, Wikipedia, social media, meta-data, ...
- Labor: Crowdsourcing and human computation
  - Games with a Purpose, Mechanical Turk, Citizen Science, crowdsourcing contests, ...

Introduction

Contributions from users increasingly central to Web

- Information: Prediction markets, ratings, opinion polls, ...
- Content: User-generated content (UGC)
  - Reviews, Q&A forums, Wikipedia, social media, meta-data, ...
- Labor: Crowdsourcing and human computation
  - Games with a Purpose, Mechanical Turk, Citizen Science, crowdsourcing contests, ...

Quality, participation, vary widely across sites

Incentives and strategic behavior

- User behavior depends on *incentives*
  - Evidence (anecdotal, formal) of self-interested users
  - 'Goodness' of system's output depends on user behavior

Incentives and strategic behavior

- User behavior depends on *incentives*
  - Evidence (anecdotal, formal) of self-interested users
  - 'Goodness' of system's output depends on user behavior
- Formal incentive design for social computing and UGC

Incentives and strategic behavior

- User behavior depends on *incentives*
  - Evidence (anecdotal, formal) of self-interested users
  - 'Goodness' of system's output depends on user behavior
- Formal incentive design for social computing and UGC
  - Agents make choices over *actions*
  - Choices of actions lead to outcomes, determine *payoffs*
  - Agents choose actions to maximize their payoffs
  - Mechanism design: Get agents to choose actions maximizing value to system

Incentives and strategic behavior

- Action choices in social computing and UGC:
  - Information: Revealing information truthfully (e.g., ratings)
  - Participation (Entry is endogenous, strategic choice!)
  - Effort: Quality of content (UGC) or output (crowdsourcing)
  - Other domain-specific choices

Incentives and strategic behavior

- Action choices in social computing and UGC:
  - Information: Revealing information truthfully (e.g., ratings)
  - Participation (Entry is endogenous, strategic choice!)
  - Effort: Quality of content (UGC) or output (crowdsourcing)
  - Other domain-specific choices
- Constructing a model: 'Capturing' strategic issues
  - Who are agents?
  - What are their costs and benefits?
  - How are agents' outputs evaluated?
  - What are 'good' outcomes?

Incentives and strategic behavior

- Action choices in social computing and UGC:
  - Information: Revealing information truthfully (e.g., ratings)
  - Participation (Entry is endogenous, strategic choice!)
  - Effort: Quality of content (UGC) or output (crowdsourcing)
  - Other domain-specific choices
- Constructing a model: 'Capturing' strategic issues
  - Who are agents?
  - What are their costs and benefits?
  - How are agents' outputs evaluated?
  - What are 'good' outcomes?

But before that...

What is social computing anyway?

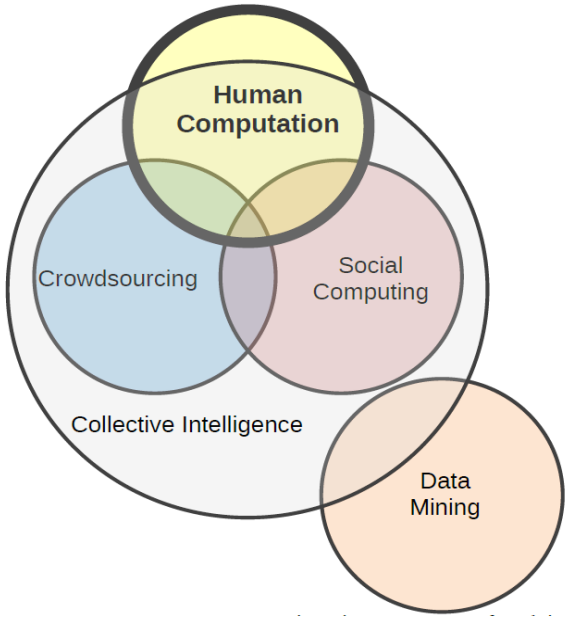
What is social computing anyway?

Social computing: "Humans in a social role where communication is mediated by technology" (Quinn and Bederson, CHI'11)

What is social computing anyway?

Social computing: "Humans in a social role where communication is mediated by technology" (Quinn and Bederson, CHI'11)

- A taxonomy of online user-contribution domains



- "Online collective action, social interaction; exchange of multimedia information, evolution of aggregate knowledge"
- Blogs, wikis, online communities, ...

Outline

- Introduction
- Eliciting information (Yiling Chen)
  - Eliciting information for events with verifiable outcomes
  - Eliciting information for events with unverifiable outcomes
- Eliciting effort and participation
  - Aspects of a model: Rewards, observability, value
  - Perfect observability: Crowdsourcing contests and other things
  - Imperfect observability: User-generated content
  - Rewarding overall contribution

Part I: Eliciting information

YILING CHEN, Harvard University

(Slides available at Yiling's webpage)

Part II: Eliciting effort and participation

ARPITA GHOSH

Cornell University

Outline

- Aspects of a model: Rewards, observability, value
- Perfect observability
  - Monetary rewards: Crowdsourcing contests, social search
  - Non-monetary attention rewards
- Imperfect observability: User-generated content
  - Private provision of public goods approach
  - Models and mechanisms: Attention, virtual points rewards
  - Introducing temporal aspects
- Unobservable output: Eliciting effort and information
- Rewarding overall contribution

Incentives for effort

- Effort as a choice in social computing
  - Review sites (Yelp, Amazon, . . .), Q&A forums (Y! Answers, StackOverflow, . . .), crowdsourcing platforms (Mechanical Turk, Citizen Science, . . .), contests (TopCoder, TaskCN), . . .

Incentives for effort

- Effort as a choice in social computing
  - Review sites (Yelp, Amazon, . . .), Q&A forums (Y! Answers, StackOverflow, . . .), crowdsourcing platforms (Mechanical Turk, Citizen Science, . . .), contests (TopCoder, TaskCN), . . .
- Quality, participation varies widely across social computing sites
- How to incentivize good outcomes?

Incentives for effort

- Effort as a choice in social computing
  - Review sites (Yelp, Amazon, . . .), Q&A forums (Y! Answers, StackOverflow, . . .), crowdsourcing platforms (Mechanical Turk, Citizen Science, . . .), contests (TopCoder, TaskCN), . . .
- Quality, participation varies widely across social computing sites
- How to incentivize good outcomes?
  - Quality: What qualities of contributions arise in equilibrium?
  - Quantity, or participation: How many, what kind of users participate in equilibrium?
  - *Endogenous* entry: Participation is typically voluntary, strategic choice

Incentives for effort

Two components to incentivizing participation and effort:

Incentives for effort

Two components to incentivizing participation and effort:

- What constitutes a reward?: Understanding contributor motivations
  - Monetary rewards
  - Social-psychological rewards: Attention [Huberman et al '09, ...], reputation and status [Beenen et al'04, ...], virtual points(!) [Nam et al '09, ...], ...

Incentives for effort

Two components to incentivizing participation and effort:

- What constitutes a reward?: Understanding contributor motivations
  - Monetary rewards
  - Social-psychological rewards: Attention [Huberman et al '09, ...], reputation and status [Beenen et al'04, ...], virtual points(!) [Nam et al '09, ...], ...
- How to **allocate** rewards?
  - Rewards are limited: *Allocate* to incentivize desirable outcomes
  - Game-theoretic framework for incentivizing effort

Game theoretic analysis for effort in social computing

- Incentivizing effort: Elements of a model
  - Agents, utilities, outcomes, 'goodness' of outcomes, ...

Game theoretic analysis for effort in social computing

- Incentivizing effort: Elements of a model
  - Agents, utilities, outcomes, 'goodness' of outcomes, ...
- How do different social computing systems differ?
  - Nature of reward and reward regimes
  - Constraints on rewards
  - Observability of (value of) agents' output
  - How output translates to value

Elements of a model: Rewards

- Different motivators lead to different rewards:
  - Money: Amazon Mechanical Turk, crowdsourcing contests (TopCoder, Innocentive, TaskCN), ...
  - Attention, reputation and status, social exchange: YouTube, Q&A forums, discussion boards, ...

Elements of a model: Rewards

- Different motivators lead to different rewards:
  - Money: Amazon Mechanical Turk, crowdsourcing contests (TopCoder, Innocentive, TaskCN), ...
  - Attention, reputation and status, social exchange: YouTube, Q&A forums, discussion boards, ...
  - *Mixed incentives*

Elements of a model: Rewards

- Different motivators lead to different rewards:
  - Money: Amazon Mechanical Turk, crowdsourcing contests (TopCoder, Innocentive, TaskCN), ...
  - Attention, reputation and status, social exchange: YouTube, Q&A forums, discussion boards, ...
  - *Mixed incentives*
- Constraints vary with nature of reward and setting:
  - Money: Arbitrary redistribution, transferable
  - Total cost as objective or constraint: Minimize cost versus budget constraint
  - Social-psychological rewards: Redistribution, transfers may be constrained

Elements of a model: Rewards

- Reward *regimes* vary with setting:
  - Bounded versus diverging rewards: Attention rewards on very popular UGC sites
  - Micro-payments (Amazon Mechanical Turk): Behavioral preferences may change wrt conventional payments

Elements of a model: Rewards

- Reward *regimes* vary with setting:
  - Bounded versus diverging rewards: Attention rewards on very popular UGC sites
  - Micro-payments (Amazon Mechanical Turk): Behavioral preferences may change wrt conventional payments
- Is total available reward exogenous or *endogenous*?
  - Exogenous: Fixed total prize money (crowdsourcing contests)
  - Endogenous: Social interaction (reward increases with increased (elicited) participation); attention rewards (higher quality content attracts larger viewership), ...

Elements of a model: Observability of output

- Perfectly observable output
  - Relative value— perfect *rank* orderings: Crowdsourcing contests (Innocentive, . . .)
  - Absolute value (less common): Output-verified piece-rate tasks (ODesk, . . .), number of recruits (social search), . . .
  - An underlying issue: How do output and value relate?

Elements of a model: Observability of output

- Perfectly observable output
  - Relative value— perfect *rank* orderings: Crowdsourcing contests (Innocentive, . . .)
  - Absolute value (less common): Output-verified piece-rate tasks (ODesk, . . .), number of recruits (social search), . . .
  - An underlying issue: How do output and value relate?
- Imperfectly observable output
  - Output value noisily reflected by ratings, votes: User-generated content (reviews, comments, answers, . . .)

Elements of a model: Observability of output

- Perfectly observable output
  - Relative value— perfect *rank* orderings: Crowdsourcing contests (Innocentive, . . .)
  - Absolute value (less common): Output-verified piece-rate tasks (ODesk, . . .), number of recruits (social search), . . .
  - An underlying issue: How do output and value relate?
- Imperfectly observable output
  - Output value noisily reflected by ratings, votes: User-generated content (reviews, comments, answers, . . .)
- Unobservable, or unverifiable, output
  - No independent evaluation of output: Crowdsourced labor (e.g. image labeling on MTurk)

Elements of a model: Effort and value

How does effort relate to *value* of contribution?

- Value mainly 'ability'-dependent: Expertise-based Q&A forums (medical, legal), ...
  - Ability: Intrinsic, or related to costly-to-acquire knowledge/skill
  - Participation-only models

Elements of a model: Effort and value

How does effort relate to *value* of contribution?

- Value mainly 'ability'-dependent: Expertise-based Q&A forums (medical, legal), ...
  - Ability: Intrinsic, or related to costly-to-acquire knowledge/skill
  - Participation-only models
- Effort-only value: Reviews, meta-data (tags, bookmarks), ...

Elements of a model: Effort and value

How does effort relate to *value* of contribution?

- Value mainly 'ability'-dependent: Expertise-based Q&A forums (medical, legal), ...
  - Ability: Intrinsic, or related to costly-to-acquire knowledge/skill
  - Participation-only models
- Effort-only value: Reviews, meta-data (tags, bookmarks), ...
- Value depends on ability and effort: Discussion forums, blogs, crowdsourcing (image classification), ...

Elements of a model: Effort and value

How does effort relate to *value* of contribution?

- Value mainly 'ability'-dependent: Expertise-based Q&A forums (medical, legal), ...
  - Ability: Intrinsic, or related to costly-to-acquire knowledge/skill
  - Participation-only models
- Effort-only value: Reviews, meta-data (tags, bookmarks), ...
- Value depends on ability and effort: Discussion forums, blogs, crowdsourcing (image classification), ...
- Associated incentive issues:
  - Adverse selection, moral hazard (Jian & MacKie Mason'13)
  - Abilities of participants who (endogenously) choose to contribute (Morgan et al '12)

Elements of a model: Effort and value

- What is the value from *multiple* contributions?

Elements of a model: Effort and value

- What is the value from *multiple* contributions?
- Participation and effort typically both strategic choices: Quality-quantity tradeoffs
  - Few very high quality or several medium-quality contributions
  - Value from highest quality contribution: Contests, some kinds of Q&A, ...
  - Value from quantity: Reviews, other Q&A (e.g., recommendations for a week in Philly), ...

Elements of a model: Effort and value

- What is the value from *multiple* contributions?
- Participation and effort typically both strategic choices: Quality-quantity tradeoffs
  - Few very high quality or several medium-quality contributions
  - Value from highest quality contribution: Contests, some kinds of Q&A, ...
  - Value from quantity: Reviews, other Q&A (e.g., recommendations for a week in Philly), ...
- Moving beyond 'vector of qualities': Value from *set* of contributions

Outline

- Aspects of a model: Rewards, observability, value
- Perfect observability
  - Monetary rewards: Crowdsourcing contests, social search
  - Non-monetary attention rewards
- Imperfect observability: User-generated content
- Unobservable output: Eliciting effort and information
- Rewarding overall contribution

Contests: Monetary rewards, perfectly observable quality

- Online crowdsourcing contests (Innocentive, TopCoder, TaskCN, . . .)

Contests: Monetary rewards, perfectly observable quality

- Online crowdsourcing contests (Innocentive, TopCoder, TaskCN, . . .)
- Contests: Pre-Internet phenomenon!
  - 1714: £20,000 contest for finding longitude at sea
  - 1829: £500 for railway engine
  - Pre-World War II: US Army contests for awarding contracts

Contests: Monetary rewards, perfectly observable quality

- Online crowdsourcing contests (Innocentive, TopCoder, TaskCN, . . .)
- Contests: Pre-Internet phenomenon!
  - 1714: £20,000 contest for finding longitude at sea
  - 1829: £500 for railway engine
  - Pre-World War II: US Army contests for awarding contracts
- An aside: *Innovation races*
  - First participant to meet some set quality bar wins
  - Contrast with contest: Fixed evaluation date, rank by quality
  - Strategic choice in models: Investment rate (Loury'79, Lee-Wilde'80, . . .)

Contests: Monetary rewards, perfectly observable quality

- Online crowdsourcing contests (Innocentive, TopCoder, TaskCN, . . .)
- Contests: Pre-Internet phenomenon!
  - 1714: £20,000 contest for finding longitude at sea
  - 1829: £500 for railway engine
  - Pre-World War II: US Army contests for awarding contracts
- An aside: *Innovation races*
  - First participant to meet some set quality bar wins
  - Contrast with contest: Fixed evaluation date, rank by quality
  - Strategic choice in models: Investment rate (Loury'79, Lee-Wilde'80, . . .)
- Netflix Prize: Features of both races and contests

Incentivizing effort in contests

- Observability: Principal can (at least) *rank* contributions perfectly
- Basic contest design problem:
  - Contestants have cost to effort
  - Principal has some total available reward budget
  - How to split reward to induce 'optimal outcomes'?

Incentivizing effort in contests

- Observability: Principal can (at least) *rank* contributions perfectly
- Basic contest design problem:
  - Contestants have cost to effort
  - Principal has some total available reward budget
  - How to split reward to induce 'optimal outcomes'?
- Designer's objective: Best, average, ... (value from output)

Incentivizing effort in contests

- Observability: Principal can (at least) *rank* contributions perfectly
- Basic contest design problem:
  - Contestants have cost to effort
  - Principal has some total available reward budget
  - How to split reward to induce 'optimal outcomes'?
- Designer's objective: Best, average, ... (value from output)
- Additional dimensions:
  - Restricting participation: Entry fees, invitation, ...
  - *Staging*: Single-stage, or multi-stage evaluation

The literature on contest design

Tournament design: Digging for Golden Carrots [Taylor'95]

- Potential participants: Pool of ex-ante identical firms
- Sponsor: Invites N firms to compete for prize P
  - Winner chosen in period T
  - Each firm pays entry cost E
  - Firm i 'trying' in period t draws $x_{i,t}$ IID at cost C
  - Prize awarded to firm $i = \arg\max_{i,1 \leq t \leq T} x_{i,t}$
- Sponsor chooses N, E
  - Free entry ($E = 0, N = N_{\max}$) is *not* optimal
  - Optimal tournament restricts entry, taxes entrants $E > 0$

Contest design: Is awarding one prize optimal anyway?

Optimal Allocation of Prizes in Contests [Moldovanu-Sela'01]

- N heterogenous contestants with differing *abilities*
  - N exogenously fixed
- Single-period game: Contestants choose *effort*
  - Cost of effort $c(e)$ increasing in e
  - Perfectly-observable effort: Highest effort wins first prize, ...
- Total prize P: Can be arbitrarily allocated amongst $1, \ldots, N$
- Designer wants to maximize *total* elicited effort
  - c linear or concave: Winner-take-all is optimal
  - c convex: Optimal contest may award multiple prizes

Contest design: Moving beyond prize-splitting

Contest *architecture* [Moldovanu-Sela'06]

- Sub-contests with winners competing, versus single contest
- Designer's objective: Maximizing total or best contribution
- Linear effort cost $c(e)$
  - Maximizing total effort: Single winner-take-all contest is optimal
  - Maximizing highest effort: Two-divisional final
- Convex costs: Many sub-contests, prizes to finalists may be optimal

Contest design: Crowdsourcing contests

- Optimal design of crowdsourcing contests [Archak-Sundararajan'09]
  - Principal's utility: Sum of top k contributions minus prize P
  - Linear effort costs, heterogenous participants
  - Risk-neutral agents: Winner-take-all contest is optimal
  - Risk-averse agents: Award multiple prizes
  - Structure of equilibrium in large N limit
- Other aspects of crowdsourcing contest design
  - Wasted effort from sunk costs [Chawla et al'11]
  - Endogenous entry [Ghosh-McAfee'12]
  - ...

Perfectly observable quality: Social computing on networks

- Information search on social network
  - Root node wants information, located somewhere in social network
  - How to incentivize search on network?

Perfectly observable quality: Social computing on networks

- Information search on social network
  - Root node wants information, located somewhere in social network
  - How to incentivize search on network?
- Monetary rewards, (first-order) perfectly observable quality

Perfectly observable quality: Social computing on networks

- Information search on social network
  - Root node wants information, located somewhere in social network
  - How to incentivize search on network?
- Monetary rewards, (first-order) perfectly observable quality
- Query-incentive networks [Kleinberg-Raghavan'05]
  - Root values query at v
  - Promises rewards to neighbors if supply answer
  - Every node can propagate query, promise reward-share
  - Size of query incentive v to obtain answer as function of answer'rarity', network structure

Social computing on networks: A real instance

- The DARPA red balloon challenge (2009)
  - 10 red balloons, distributed across US
  - First team to correctly locate all balloons wins $40,000
- Incentives challenges:
  - Incentivize agents to participate
  - Each participant should want to incentivize others to participate
- Won by MIT team in under 9 hours
  - Recruited ≈4400 participants in under 36 hours

Social computing on networks: A real instance

'Recursive incentive scheme' used by winning team [Pickard et al'11]

- Exponential reward structure, decreasing down from 'finders' to root:
  - Never runs a deficit: Respects total budget
  - Incentivizes further recruitment
  - Does not create incentives to bypass 'inviters'

Social computing on networks: A real instance

'Recursive incentive scheme' used by winning team [Pickard et al'11]

- Exponential reward structure, decreasing down from 'finders' to root:
  - Never runs a deficit: Respects total budget
  - Incentivizes further recruitment
  - Does not create incentives to bypass 'inviters'
- *Does* provide incentives for false-name attacks: Output is not quite "perfectly observable"!
  - Sybil-proof mechanisms [Babaioff et al'12]

Perfectly observable quality: Non-monetary rewards

- Attention rewards in crowdsourced *content*
  - Q&A forums (Quora, StackOverflow, . . .)
  - Asker can supply rank-ordering of answers

Perfectly observable quality: Non-monetary rewards

- Attention rewards in crowdsourced *content*
  - Q&A forums (Quora, StackOverflow, ...)
  - Asker can supply rank-ordering of answers
- Total reward cannot be (reasonably) arbitrarily redistributed
  - Attention at position $i+1$ is subset of attention at i
  - Maximum available attention at position i: A_i
  - Constraint: Choose reward $a_i \leq A_i$
  - Note contrast with sum constraint from monetary rewards ($\sum_{i=1}^{n} a_i \leq B$)

Optimal mechanisms for attention rewards

- UGC site can suppress content ($a_i < A_i$): Eliminate; display less prominently...
  - Payoff to poor quality falls, but so does participation
  - What $a_i \leq A_i$ lead to 'best' outcomes?

Optimal mechanisms for attention rewards

- UGC site can suppress content ($a_i < A_i$): Eliminate; display less prominently...
  - Payoff to poor quality falls, but so does participation
  - What $a_i \leq A_i$ lead to 'best' outcomes?

- Full reward to all but lowest possible rank optimizes entire *equilibrium distribution* of qualities: [Ghosh-McAfee'12]

$$a_i = A_i, i = 1, \ldots, n-1; \quad a_n = \min(A_n, c(0))$$

  - Optimal reward for lowest possible rank depends on cost of producing lowest quality
  - Reward structure optimal for *any* increasing function of qualities (accounting for endogenous participation)

Outline

- Aspects of a model: Rewards, observability, value
- Perfect observability
- Imperfect observability: User-generated content
  - Private provision of public goods approach
  - Models and mechanisms: Attention, virtual points rewards
  - Introducing temporal aspects
- Unobservable output: Eliciting effort and information
- Rewarding overall contribution

Imperfectly observable quality

- Imperfectly observable quality:
  - Agents' effort does not perfectly map into output
  - Effort maps perfectly to output, but system cannot observe quality perfectly: Noisy observations

Imperfectly observable quality

- Imperfectly observable quality:
  - Agents' effort does not perfectly map into output
  - Effort maps perfectly to output, but system cannot observe quality perfectly: Noisy observations
- An aside: Who makes observations of output?
  - Disinterested observers: Raters with no agenda (Typical assumption in literature so far)
  - Strategic raters (such as competitors): We'll return to this briefly later

Imperfectly observable quality: User-generated content (UGC)

- UGC: Information contributed by users with no *direct*, *extrinsic* compensation [MacKie-Mason'09]
  - Reviews (Amazon, Yelp, TripAdvisor, ...)
  - Knowledge-sharing forums (Quora, StackOverflow, Y!A, ...)
  - Comments (Slashdot, News, ...)
  - Social media (Blogs, YouTube, Flickr, ...)
  - Metadata: Tags, bookmarks (del.icio.us, ...)
- No single principal who evaluates quality: User-contributed ratings
  - Infeasible: Scale of contributions

Imperfectly observable quality: User-generated content (UGC)

- UGC: Information contributed by users with no *direct*, *extrinsic* compensation [MacKie-Mason'09]
  - Reviews (Amazon, Yelp, TripAdvisor, ...)
  - Knowledge-sharing forums (Quora, StackOverflow, Y!A, ...)
  - Comments (Slashdot, News, ...)
  - Social media (Blogs, YouTube, Flickr, ...)
  - Metadata: Tags, bookmarks (del.icio.us, ...)
- No single principal who evaluates quality: User-contributed ratings
  - Infeasible: Scale of contributions
  - Value of content not determined by any single agent

Incentives in user-generated content

- Key issues in UGC
  - Quantity: Inducing adequate participation
  - Quality: Contributions are not homogenous (unlike monetary donations)
  - Evaluation: Quality not easy or cheap to evaluate
  - Spam: Detect and remove rather than disincentivize (not this talk)
- How to incentivize high quality and participation? Modeling incentives in UGC

Incentives in user-generated content

- Key issues in UGC
  - Quantity: Inducing adequate participation
  - Quality: Contributions are not homogenous (unlike monetary donations)
  - Evaluation: Quality not easy or cheap to evaluate
  - Spam: Detect and remove rather than disincentivize (not this talk)
- How to incentivize high quality and participation? Modeling incentives in UGC
  - What is quality?
  - Agents and strategies, evaluation, ...
  - Objectives, metrics

Incentives in UGC: A public-goods perspective

Incentivizing UGC: Private provision of public goods [Jian-MacKie Mason'09,'13]

- Distinctions from charitable giving:
  - Non-homogeneity of contributions: Content qualities unequal
  - No side payments, or direct transfers
- Technology-reliant incentives:
  - Functionality in exchange for content (bookmarking, photo storage, ...)
  - Provide social motivators (interaction, reputation, ...)
  - *Exclusion* mechanisms: Block or limit access to content

Incentives in UGC: A public-goods perspective

- *Exclusion mechanisms*: Content need not remain pure public good
  - Limit usage based on contribution level: Glassdoor, P2P, ...
  - Cost-sharing literature: Raising money versus effort, for public good of known value, ...
- Effectiveness of minimum-threshold mechanisms
  - Evidence from lab experiments [Swope'02, Kocher'05, Croson'08]
  - Game-theoretic analysis of efficiency [Wash-MacKie-Mason'08, Jian'10]

Incentivizing high-quality UGC: Attention rewards

- A game-theoretic model with attention rewards [Ghosh-McAfee'11]
  - Attention motivates contributors: Empirical studies
  - Model, analysis, agnostic to *why* users like attention
- *Quality* is a probability $q \in [0, 1]$
  - q: Probability viewer answers yes to "Is content useful?"
- Contributors: Cost to quality, benefit from attention
  - Strategies: Quality *and* participation
- Evaluation: Non-strategic viewers rate according to q

Incentivizing high-quality UGC: Attention rewards

- Metric: Asymptotic (diverging-attention limit) equilibrium participation, quality of contributions
- Asmptotically optimal quality, high participation achievable with simple elimination mechanism
- Incentives in rank-order UGC mechanisms [Ghosh-Hummel'11]
  - Also incentivize near-optimal quality, high participation: Need adequate separation between ranks; zero attention to very low ranks
  - Equilibrium quality 'dominates' that in proportional mechanism

User-generated content: Virtual points rewards

- Virtual points motivate users [Nam et al'09, . . .]
  - Online Q&A forums (Y! Answers, Naver, . . .),
  - Non-diverging, arbitrarily allocatable
- Many online Q&A forums use *best-contribution* mechanisms
  - Winner gets p_B, everyone else gets p_C
- Can $(p_B; p_C)$ structure *implement* optimal outcomes? [Ghosh-Hummel'12]
  - Yes: When contribution's value largely determined by expertise
  - When value depends on expertise and effort: Only possibly with noisy rankings!

Imperfectly observable quality: Strategic rating

- System cannot directly observe qualities: Relies on user ratings
- What if voters are *strategic*?
  - Voters also contributors of content
  - Compete with other contributors for high rankings

Imperfectly observable quality: Strategic rating

- System cannot directly observe qualities: Relies on user ratings
- What if voters are *strategic*?
  - Voters also contributors of content
  - Compete with other contributors for high rankings
- Voting for the voters [Alon et al'11]
  - Approval voting: Every voter is also a candidate
  - Strategyproof, approximately optimal mechanisms for selecting k-best subset

Imperfectly observable quality: Strategic rating

- System cannot directly observe qualities: Relies on user ratings
- What if voters are *strategic*?
  - Voters also contributors of content
  - Compete with other contributors for high rankings
- Voting for the voters [Alon et al'11]
  - Approval voting: Every voter is also a candidate
  - Strategyproof, approximately optimal mechanisms for selecting k-best subset
- Incentivizing *contribution*, in presence of strategic raters: Need for models, mechanisms

Temporal issues

- Literature discussed so far (largely) did not model time
  - Simultaneous contributions by agents
  - No temporal aspect to evaluating output

Temporal issues

- Literature discussed so far (largely) did not model time
  - Simultaneous contributions by agents
  - No temporal aspect to evaluating output

- Time and input: Eliciting timely contributions in Q&A forums [Jain et al'09]
  - Action choice: Time of contribution (no quality dimension)
  - Reward allocation rules and efficiency of equilibria

Temporal issues

- Literature discussed so far (largely) did not model time
  - Simultaneous contributions by agents
  - No temporal aspect to evaluating output
- Time and input: Eliciting timely contributions in Q&A forums [Jain et al'09]
  - Action choice: Time of contribution (no quality dimension)
  - Reward allocation rules and efficiency of equilibria
- Time and output: Incentives when qualities are *learnt* [Ghosh-Hummel'13]
  - (Simultaneous) contributions rated over time; attention reward from each rater
  - Multi-armed bandit problem with *endogenous arms*

Outline

- Aspects of a model: Rewards, observability, value
- Perfect observability
- Imperfect observability: User-generated content
- Unobservable output: Eliciting effort and information
- Rewarding overall contribution

Unobservable quality

Eliciting effort when quality is unobservable [Dasgupta-Ghosh'13]

- Crowdsourced judgement: Image labeling and identification, content rating, peer grading, ...

Unobservable quality

Eliciting effort when quality is unobservable [Dasgupta-Ghosh'13]

- Crowdsourced judgement: Image labeling and identification, content rating, peer grading, ...
  - Unobservable ground truth
  - *Effort*-dependent accuracy
  - Information elicitation, with *endogenous* proficiency

Unobservable quality

Eliciting effort when quality is unobservable [Dasgupta-Ghosh'13]

- Crowdsourced judgement: Image labeling and identification, content rating, peer grading, ...
  - Unobservable ground truth
  - *Effort*-dependent accuracy
  - Information elicitation, with *endogenous* proficiency
- Main result: Mechanism where maximum effort-truthful reporting is *highest-payoff* equilibrium (No task-specific collusions)
  - Reward for agreement, but also
  - Subtract statistic term penalizing *blind* agreement: Designed so agents receive zero payoff without effort

Outline

- Aspects of a model: Rewards, observability, value
- Perfect observability
- Imperfect observability: User-generated content
- Unobservable output: Eliciting effort and information
- Rewarding overall contribution

Moving beyond single tasks: Incentivizing overall contribution

- So far: Models, incentives for *single* action/contribution/task

Moving beyond single tasks: Incentivizing overall contribution

- So far: Models, incentives for *single* action/contribution/task
- Rewarding contributors for *overall* identity:
  - Site-level accomplishments based on cumulative contribution: Badges, leaderboards, reputations...

Moving beyond single tasks: Incentivizing overall contribution

- So far: Models, incentives for *single* action/contribution/task
- Rewarding contributors for *overall* identity:
  - Site-level accomplishments based on cumulative contribution: Badges, leaderboards, reputations...
- Rewards valued by users: Increased engagement
  - Reputation: Value online and offline (StackOverflow, ...)
  - Badges: Formal inference from data [Anderson et al'13]
  - Anecdotal: Online discussion boards for Amazon Top-Reviewer list, Y! Answers Top-Contributor badge

Rewarding for overall contribution

- *Gamification*: Badges, leaderboard positions, ...
- Different gamification *designs* online:
  - StackOverflow, Foursquare, ...: Badges for activity milestones ('Absolute' badges)
  - Y!Answers, Quora, Tripadvisor: Top Contributors ('Competitive' badges)
  - Amazon: Top Reviewer List (Rank-order rewards)
- Rewards valued by users, but require costly effort: Gamification induces *mechanism*
- How to 'gamify' to incentivize participation, effort?

Some questions

- What incentives are created by absolute badges, and by competitive badges?
- How do these incentives compare?
- Should competitive badges be awarded to fixed number or fraction of participants?
- Should multiple absolute badges be awarded, and if yes, 'how'?
- How should rank-based rewards be designed?
- What if value from winning depends on other winners?

Some answers

Incentives and badge designs [Easley-Ghosh'13]

- Absolute standards mechanism $\mathcal{M}_\alpha$: Badges for output α
- Relative standards mechanism $\mathcal{M}_\rho$: Badges to top ρ contributors
- Equilibrium analysis of incentives:
  - $\mathcal{M}_\rho$: Equilibria exist if announce fixed number of badges, rather than fraction
  - Partial equivalence between $\mathcal{M}_\alpha$ and $\mathcal{M}_\rho^p$; includes optimal effort point
  - $\mathcal{M}_\alpha$ less 'robust' than $\mathcal{M}_\rho$ to uncertainty about population
- Externalities: Endogenously determined badge valuations

Some answers: II

- Multiple 'absolute standards' badges [Anderson et al'13]
  - Single action: What levels to reward to sustain effort?
  - Multiple actions: How to steer user behavior?

- Model: Multi-dimensional space of action types
  - Users incur cost to actions differing from 'preferred' distribution over actions
  - Time-discounted value to winning badges
  - Effort choice depends significantly on badge 'placement'
  - 'Spread out' multiple badges with roughly equal values

Games of status

- Reward mechanisms when users only care about status [Dubey-Geanakoplos'09]:
  - N agents each receive a grade
  - Utility from status:# with strictly lower - # with strictly higher grade
  - Output is noisy perturbation of input effort
  - What scheme incentivizes maximum effort from every agent?
- Main results:
  - Optimal grading scheme not 'perfectly fine': Coarsening improves effort
  - Absolute grading dominates grading on a curve
- Models, mechanisms: Status and social computing

Overall contribution: Temporal issues

- Frequency of contribution

Overall contribution: Temporal issues

- Frequency of contribution
  - Game-theoretic model for education forums (Piazza, MOOCs) [Ghosh-Kleinberg'13])
  - Strategic choice: *Rate* of participation
  - Cost to rate; benefit from timely contribution

Overall contribution: Temporal issues

- Frequency of contribution
  - Game-theoretic model for education forums (Piazza, MOOCs) [Ghosh-Kleinberg'13])
  - Strategic choice: *Rate* of participation
  - Cost to rate; benefit from timely contribution
- Reputation as reward for overall contribution:
  - Incentivizing quality: Reputation as input to learning algorithm
  - Participation: Update rules to elicit frequent contribution

Overall contribution: Temporal issues

- Frequency of contribution
  - Game-theoretic model for education forums (Piazza, MOOCs) [Ghosh-Kleinberg'13])
  - Strategic choice: *Rate* of participation
  - Cost to rate; benefit from timely contribution
- Reputation as reward for overall contribution:
  - Incentivizing quality: Reputation as input to learning algorithm
  - Participation: Update rules to elicit frequent contribution
- *Sustaining contribution*:
  - Model for contributor interest over time
  - Mechanisms to incentivize sustained contribution

(More) open directions

- More nuanced models of quality, output:
  - Vertical and horizontal differentiation [MacKie Mason'09], diversity
  - Modeling value from set of contributions
- Incentives for production with strategic ratings
- Overall contributor reward design
  - So far: Simple models, reward structures
  - Leaderboard design: Unequal rewards to winners

(Even more) open directions

[Ghosh'13: *Game Theory and Incentives in Human Computation*]

- Different participant roles (contribution, moderation, ...)
  - Interaction between role-specific incentives
  - Endogenous ability-based selection into roles

(Even more) open directions

[Ghosh'13: *Game Theory and Incentives in Human Computation*]

- Different participant roles (contribution, moderation, . . .)
  - Interaction between role-specific incentives
  - Endogenous ability-based selection into roles
- Mixed incentives:
  - How do users cumulatively value, tradeoff differing incentives?
  - Models, mechanism design

(Even more) open directions

[Ghosh'13: *Game Theory and Incentives in Human Computation*]

- Different participant roles (contribution, moderation, . . .)
  - Interaction between role-specific incentives
  - Endogenous ability-based selection into roles
- Mixed incentives:
  - How do users cumulatively value, tradeoff differing incentives?
  - Models, mechanism design
- *Behavioral economics*: User valuations of social-pyschological rewards
  - 'Shape' of reward functions: Marginal benefits (attention, . . .)
  - Value from *set* of rewards
  - How do rewards retain value over time?