Emergence of Equilibria from Individual Strategies in Online Content Diffusion

Eitan Altman*, Francesco De Pellegrini◇, Rachid El-Azouzi†, Daniele Miorandi◇ and Tania Jimenez†

***Abstract*—Social scientists have observed that human behavior in society can often be modeled as corresponding to a threshold type policy. A new behavior would propagate by a procedure in which an individual adopts the new behavior if the fraction of his neighbors or friends having adopted the new behavior exceeds some threshold. In this paper we study the question of whether the emergence of threshold policies may be modeled as a result of some rational process which would describe the behavior of non-cooperative rational members of some social network. We focus on situations in which individuals take the decision whether to access or not some content, based on the number of views that the content has. Our analysis aims at understanding not only the behavior of individuals, but also the way in which information about the quality of a given content can be deduced from view counts when only part of the viewers that access the content are informed about its quality. In this paper we present a game formulation for the behavior of individuals using a meanfield model: the number of individuals is approximated by a continuum of atomless players and for which the Wardrop equilibrium is the solution concept. We derive conditions on the problem's parameters that result indeed in the emergence of threshold equilibria policies. But we also identify some parameters in which other structures are obtained for the equilibrium behavior of individuals.**

***Index Terms*—User-generated content, Complex Systems, Video popularity, Game theory, Wardrop equilibria**

I. Introduction

Online media constitute currently the largest share of Internet traffic. A large part of such traffic is generated by platforms that deliver user-generated content (UGC). This includes, among the other ones, YouTube and Vimeo for videos, Flickr and Instagram for images and all social networking platforms.

Among such services, a prominent role is played by YouTube. Founded in 2005 by Chad Hurley, Steve Chen and Jawed Karim and acquired in 2006 by Google, YouTube scored in 2011 more than 1 trillion views (or, alternatively, an average of 140 video views for every person on Earth), with more than 3 billion hours of video watched every month and 72 hours of video uploaded every minute by YouTube's users[1].

Of course, not all videos posted on YouTube are equal. The key aspect is their "popularity", broadly defined as the number of views they score (also referred to as *viewcount*). This is relevant from a twofold perspective. On the one hand, more popular content generates more traffic, so understanding popularity has a direct impact on caching and replication strategy that the provider should adopt. On the other one, popularity has a direct economic impact. Indeed, popularity or viewcount are often directly related to click-through rates of linked advertisements, which constitute the basis of the YouTube's business model.

Recently, a number of researchers have analysed the evolution of the popularity of online media content [1], [2], [3], [4], [5], [6], with the aim of developing models for early-stage prediction of future popularity [7].

Such studies have highlighted a number of phenomena that are typical of UGC delivery. This includes the fact that a significant share of content gets basically no views [6], as well as the fact that popularity may see some bursts, when content "goes viral" [4]. Also, in [7] the authors demonstrate that after an initial phase, in which contents gain popularity through advertisement and other marketing tools, the platform mechanisms to induce users to access contents (re-ranking mechanisms) are main drivers of popularity.

In this paper, we address such phenomena, by developing a model, based on game theoretical concepts and tools, for understanding how user's behaviour drives the evolution of popularity of a given content. The work is based on rational decision-making assumptions, whereby the users have to decide whether to see a given content or not. This configures as a game, where users seek to maximize some expected utility based on their "perception" of the quality of the content[2] and on viewcount. However, users suffer also a cost for accessing contents of bad quality, i.e., waste of time and possibly bandwidth, batteries, etc. In particular, in the decision process the viewcount is used as a noisy estimator of the quality of a content. Interestingly, this context resembles closely the situation in the economic domain, where customers of a firm which are uninformed do infer the quality of products from the length of the queue they encounter upon requesting firm's goods to purchase [8].

Extensive advertising and marketing campaigns can be used to push the viewcount of a given content up. And in the decision making process users do not know whether the viewcount has been "pushed" by such means. Also, the decisions made by different users influence the viewcount and consequently the decisions made by other users, a process which suits well the usage of game theoretical machinery.

Specifically, we describe the conditions for the adoption of

*INRIA B.P.93, 2004 Route des Lucioles, 06902 Sophia-Antipolis, Cedex, FRANCE, email: eitan.altman@sophia.inria.fr

◇CREATE-NET, via Alla Cascata 56 c, 38100 Trento, ITALY, email: {fdepellegrini,dmiorandi}@create-net.org

†CERI/LIA, University of Avignon, 74 rue Louis Pasteur, 84029 AVIGNON Cedex 1, email: {rachid.elazouzi, tania.altman}@univ-avignon.fr

[1] http://www.youtube.com/t/press_statistics/

[2] This may come, e.g., from the name of user who posted the content.

common behaviors in online content access. This is inspired by findings in social science [9], [10], [11]: results there show that emerging behaviours would propagate by a procedure in which an individual adopts a novel behavior if the fraction of neighbors or friends having adopted the same behavior exceeds some threshold. In our context, the threshold would be expressed in terms of viewcount or related metric.

In the sense of game theory, users of online media represent non-cooperative rational players connected through some social tie, e.g., being users of the same UGC platform. Since we consider systems composed by a very large number of users, the customary tool to study the user behaviour is that of Wardrop equilibria [12]. In particular, we have found a number of conditions for which such equilibria exist and can be characterized analytically. Explicit conditions were found for content to stay at zero views or to become so popular that it is makes sense for all users to access it the sooner the better.

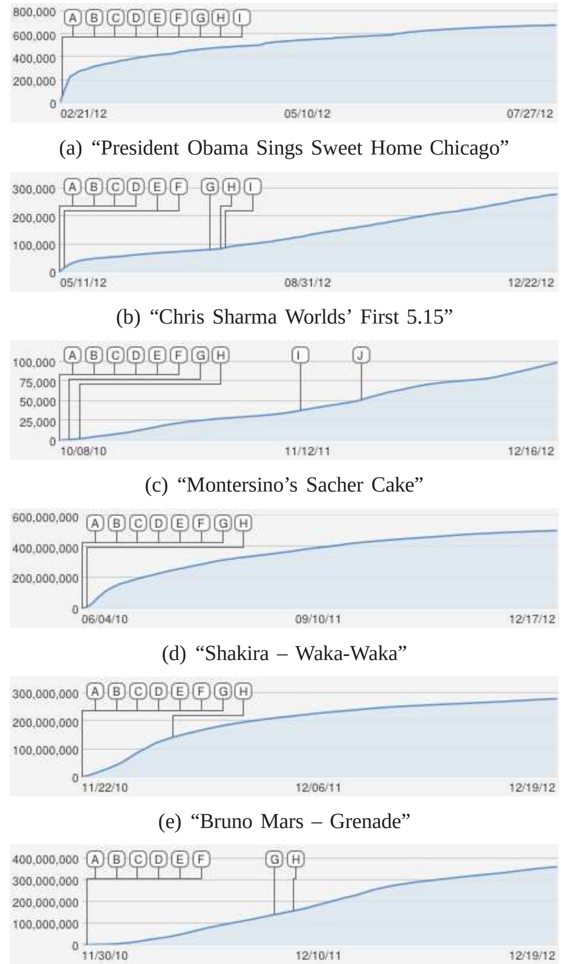
Furthermore, we identify, for the general case, conditions under which players tend to accrue around a common strategy depending on initial conditions. This is due to the existence of a continuum of equilibria: the system will settle at any point very much depending on initial conditions imposed, for instance, by a set of forerunners which cause significant changes of the content popularity. Such conditions were identified in early works such as [13] in other contexts: there, the authors applied threshold type Nash equilibrium strategies in which one purchases priority if and only if upon arrival the queue size is larger than some threshold value. Key motivation in [13] is predictability and control of purchase priority. What motivates this work is predictability and control of online content access.

Novel contribution: in this paper, we move away from the classical analysis of social networks in the spirit of [7], [4], [5], [1]: instead, we provide a first analysis based on games. The aim of this paper is to provide a novel perspective where contents compete to gain popularity and are subject to the effect of user's choice. To the best of the authors' knowledge, this is the first attempt so far to describe content popularity in UGC systems using game theoretical tools.

The remainder of the work is organized as follows. In Sec. II we introduce the system model and the notation used throughout the paper. Results for the case when plain viewcount is used to make decisions are presented in Sec. III. When decisions account also for a large increasing trend of content popularity, i.e., looking for 'hot' content, the dynamics of the game becomes different. This case is analysed in Sec. IV. In Sec. V we analyze the joint effect when both the viewcount and its trend are both relevant to the user. Finally, in Sec. VI we model the effect of side information when users have some measure of future content dynamics. Sec. VII reviews the related work and Sec. VIII concludes the paper highlighting directions for expanding the current reach of the work.

II. System Model

We consider contents made available to a user by means of YouTube or a similar platform. We denote by τ the lifetime of a content, i.e., the time horizon during which the content bears some interest. In general, such horizon differs depending on the type of content: it can be typically of the order of weeks to months for YouTube videos or a few days for news [7]. A possible extension to the case of variable time horizon is the addressed in Sec. IV.



(a) "President Obama Sings Sweet Home Chicago"

(b) "Chris Sharma Worlds' First 5.15"

(c) "Montersino's Sacher Cake"

(d) "Shakira – Waka-Waka"

(e) "Bruno Mars – Grenade"

(f) "Adele – Rolling in the deep"

Fig. 1: Dynamics of the viewcount for six sample videos: the push dynamics can be identified with the first part of the dynamics, where labels identify some actions that are significant for the diffusion of the video; observe for cases a, b and c how a linear dynamics takes over in the last part of the dynamics. The labels tagging the first part of the dynamics mention specific events that identify the diffusion of the content on specific platforms or channels.

We denote by $X(t)$ the viewcount attained by a given content θ at time t seconds after it has been posted, for $0 \le t \le \tau$.

As in standard UGC platforms, there are two mechanisms that coexist and can jointly increase the viewcount:

- *push*: the content provider exploits some preferential channels (including paid advertisement either directly on the UGC system or via social networking platforms) to make users aware of the content and to induce them to access it. We call *push users* the users that access the

content as a reaction to the push mechanism.
- *pull*: users find about the content through standard search and decide to access it based on the belief that the content is relevant for them. We call users accessing a content through the pull mechanism *pull users*.

In practice, many YouTube videos are subject to the push and the pull mechanisms described above such as the examples that we reported in Fig. 1. For instance, Fig. 1a, shows the dynamics of a popular video with viewcount $X \geq 675000$. The YouTube statistics associated with the video describe explicitly a series of events happening in the first part of the dynamics of X. For instance, the event B that appears around 02/12/2012, is precisely the event `‘‘First embedded on: plus.google.com’’` which indeed configures as a push towards a social network platform. After the initial push, such events vanish, and the rest of the dynamics appears ascribed mostly to the pull mechanism defined above, with a linear increase in the viewcount.

Also, some of the reported videos are representative of a specific class of online contents, which are those we will be dealing with in the rest of the paper. We can refer to those as the contents that comply to the *exponential-linear* model, for the sake of brevity. In particular, many such contents appear to obey to the following dynamics: after an initial exponential growth, the increase of the viewcount becomes linear. The way to interpret such a behavior can be traced to the notion of push and pull mechanisms described above: the exponential growth corresponds to actions through which the source distributes the content within a basin of target push viewers. When such basin is finite and small with respect to the content diffusion dynamics, the viewcount dynamics experiences a saturation effect which takes over after an initial phase. However, at that stage, the access to the content is due to pull users that come across the content browsing online: they do so at random from a very large basin, so that the access rate, i.e., the viewcount increase rate, is linear. These combined effects are visible in the case of the first two videos, i.e. Fig 1a and Fig 1b. In the case of the first video, the saturation effect is well visible, whereas in the case of the second one the linear increase following the saturation is dominating. The example in Fig 1c is a case where all the dynamics is linear with good approximation: as it will be clear in the following, in the exponential-linear model this case is represented when either the basin of push users is large or when the rate at which contents are pushed is small.

Remark 1: Not all videos will diffuse according to the proposed exponential-linear model. For instance, there exist cases when the initial viewcount dynamics displays a characteristic sigmoid shape. We reported in Fig 1d,e,f the viewcount dynamics for three popular music videos: in those cases the dynamics resembles the logistic curve associated to the spread of epidemics. We can ascribe such similarity to the presence of a positive feedback in the push mechanism, e.g., those who access the content have some mean to recommend the content for others to access it, through targeted recommendation or similar mechanism. When a social network is present, this may happen due to the push of the content into the neighborhood of those who view the content. A similar and perhaps more powerful feedback effect can happen between different channels on the same platform, e.g., YouTube channels, and across different platforms through the recommendation list that is presented to the platform users.
This also qualifies the type of exponential-linear dynamics that we consider as those for which this type of feedback does not play a significant role. In particular, in the case of Fig 1a, the content is of interest at the national scale in the US, and the viewers are likely driven to the content by general search criteria (e.g., typing in a search engine). Also, in the case of Fig 1c, the viewers are likely those who browse for some specific recipe, whereas in the case of Fig 1b viewers are interested in a niche sport, where the event is known within the reference community. In all such cases we see that the linear part of the dynamics takes over and becomes dominant.

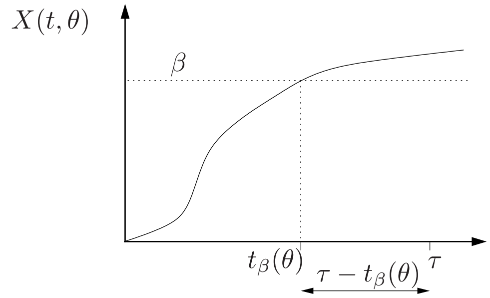


Fig. 2: The reward or the cost of content θ for a tagged user is represented by the time during which the content can be accessed, i.e., when viewcount is larger than threshold β.

Game model

In our model, we are interested in the uptake of the pull users. Pull users interested in the given content do not know in advance its quality. They may discover it during interval $[0, \tau]$ at random. Their estimation of the interest/potential quality is based on the viewcount X. In the simplest case, contents with higher viewcount are more likely to be accessed.

We define by $X_{ps}(t)$ the number of push users accessing the content up to time t as a reaction to the push mechanism and, analogously, by $X_{pu}(t)$ the number of those accessing it through the pull mechanism. Clearly, $X(t) = X_{ps}(t) + X_{pu}(t)$.

Users have beliefs about the quality of the content. We denote by π_G the belief that a given content is good (i.e., of interest or anyway worth accessing) and, conversely, by $\pi_B = 1 - \pi_G$ the belief that the content is bad. We denote by $\boldsymbol{\pi} = (\pi_G, \pi_B)$ the corresponding distribution. Stating $\pi_G = 0.75$ means that a user believes that every 4 similar contents she would get 3 good ones and 1 bad one.

The content access configures as a game where we define *players*, *strategies* and *utilities*. *Players:* the *players* are pull users: based on their belief $\boldsymbol{\pi}$, they may access the content θ or not.

Strategies: they access θ when the viewcount is above a certain threshold, i.e., $X(t) \geq \beta \geq 0$.[3] Hence, the *strategy* for a certain user is the viewcount threshold $\beta \geq 0$. Of course, all other players also adopt their own strategy with respect to θ and we denote $\boldsymbol{\alpha}$ the vector of strategies of all remaining users: $\boldsymbol{\alpha}$ is a vector of viewcount thresholds for all other users.

Utilities: users face either a *cost* C or a *reward* R for playing strategy β: the cost and the reward is the fraction of lifetime when the content is in the viewcount range, i.e, when they are willing to access it. The rationale to define this cost/reward is the following. Let a good content be worth one unit reward, and a bad content worth a unit cost. The user may hit several similar contents at random over time. If they are good, the fraction of those actually accessed will be proportional to $1 - \frac{t_\beta}{\tau}$, where we define $t_\beta = \min\{t \,|\, \beta = X(t)\}$, i.e., t_β is the smallest instant when the threshold is achieved. That also is going to be the long term reward, or the cost, for accessing similar online contents. Formally,

$$R(\boldsymbol{\alpha}, \beta, G) = (\tau - t_\beta(G))^+, \quad C(\boldsymbol{\alpha}, \beta, B) = (\tau - t_\beta(B))^+$$

Finally, based on their belief $\boldsymbol{\pi}$, players expect a utility when playing β that amounts to

$$U(\boldsymbol{\alpha}, \beta) = \pi_G R(\boldsymbol{\alpha}, \beta, G) - \pi_B C(\boldsymbol{\alpha}, \beta, B)$$

According to the above expression, the cost and the reward are a function of the interval when the content is above the threshold, i.e., when the users can benefit from it, and depends on the other players strategy. Furthermore, the action taken by players depends on their belief on the quality of the content.

In the following we will investigate *symmetric equilibria*, i.e., equilibria for which all users play $\alpha \geq 0$. We can hence adopt a simplified scalar notation and define $t_\alpha = \min\{t | \alpha = X(t)\}$.

Let a tagged user playing β when all the remaining users use α: we make the assumption that Wardrop conditions holds. Namely, for a large number of users any unilateral deviation of a single user does not affect the utilities of other users. I.e., deviations due to a single user action are negligible. Wardrop equilibria are much easier to compute than the Nash equilibrium; however, Wardrop is a good approximation for the latter, as in [14].[4].

The tagged user expects to gain a certain reward $R(\alpha, \beta, G)$, for a good content and expects to suffer a cost $C(\alpha, \beta, B)$ when the content is bad: under which conditions α is the best response to itself, namely $\beta^*(\alpha)$? We answer to this question in the next sections under different knowledge of the viewcount dynamics available to users.

Before we introduce our analysis, we recall that the utility function has the following expression for $\beta \geq \beta_{\tau,B}$

$$U(\alpha, \beta) = \begin{cases} 0 & \text{if} \quad \beta \geq \beta_{\tau,G} \\ \pi_G(\tau - t_\beta(G)) & \text{if} \quad \beta_{\tau,B} \leq \beta \leq \beta_{\tau,G} \end{cases}$$

where $\beta_{\tau,\theta}$ is solution of the following equation

$$t_{\beta_\tau}(\theta) = \tau \tag{1}$$

We observe that the utility function U is nonincreasing for $\beta \geq \beta_{\tau,B}$. However the best response $\beta^*(\alpha)$ can be found only in the interval $[0, \beta_{\tau,B}]$. As a result we restrict our analysis to case when $\beta \leq \beta_{\tau,B}$ in which the utility function can be expressed as

$$U(\alpha, \beta) = \pi_G(\tau - t_\beta(G)) - \pi_B(\tau - t_\beta(B))$$

III. Plain Viewcount

The basic model that we introduce in this section is based on the assumption that pull users rely on the number of hits of the contents to judge if it is worth to access it or not, i.e., they judge based on how many users accessed it. Thus, they play based on the dynamics. We hence specialize our analysis to two cases.

A. Linear case

First, we examine the case when the process of diffusion of contents is linear. This is the case when the time scale of the content diffusion is very large compared to the pool of potential users. A mechanism that that is able generate such a dynamics is the combined effect of an advertisement which is broadcasted to a very large pool of viewers, e.g., covering newspapers or other general audience media, and people so made aware of the existence of the content who decide to access the content with some random delay thereafter.

Thus, we let $X_{ps}(t, \theta) = \lambda_{ps} t \cdot \mathbb{1}(t)$ where $\mathbb{1}(t)$ is the unitary step function, and $X_{pu}(t, \theta) = \lambda_{pu}(t - t_\alpha) \cdot \mathbb{1}(t - t_\alpha)$[5].

Observe that in this case $\lambda_{ps} = \lambda_{ps}(\theta)$, whereas λ_{pu} is independent of θ. In fact, we assume pull users judge based on viewcount only [8]. However, we assume that $\lambda_{ps}(G) \geq \lambda_{ps}(B)$.

Lemma 1: In the linear case, under the assumption $\lambda_{pu}(G) \geq \lambda_{pu}(B)$, it holds

i. if $\frac{\pi_G}{\lambda_{ps}(G)} \geq \frac{\pi_B}{\lambda_{ps}(B)}$, then $\beta^*(\alpha) = 0$.
ii. if $\frac{\pi_G}{\lambda_{ps}(G)} \leq \frac{\pi_B}{\lambda_{ps}(B)}$ but $\frac{\pi_G}{\lambda_{ps}(G)+\lambda_{pu}} \geq \frac{\pi_B}{\lambda_{ps}(B)+\lambda_{pu}}$, then $\beta^*(\alpha) = \alpha$
iii. if $\frac{\pi_G}{\lambda_{ps}(G)} \leq \frac{\pi_B}{\lambda_{ps}(B)}$ but $\frac{\pi_G}{\lambda_{ps}(G)+\lambda_{pu}} < \frac{\pi_B}{\lambda_{ps}(B)+\lambda_{pu}}$, then $\beta^*(\alpha) = \beta_{\tau,B}$

Proof: We need to distinguish two cases, namely $\alpha \geq \beta$ and $\alpha \leq \beta$, determine the best response for each case, and then by comparison choose the best response $\beta^* = \beta^*(\alpha)$. The expression for the utility in the two cases follows.

If $\alpha \geq \beta$, then $X(t, \theta) = X_{ps}(t, \theta)$ for $0 \leq t \leq t_\beta$. Thus, we can write simply

$$t_\alpha = \frac{\alpha}{\lambda_{ps}(\theta)}, \quad t_\beta = \frac{\beta}{\lambda_{ps}(\theta)}$$

and the expression for the utility

$$U(\alpha, \beta) = \tau(\pi_G - \pi_B) - \beta\Big(\frac{\pi_G}{\lambda_{ps}(G)} - \frac{\pi_B}{\lambda_{ps}(B)}\Big) \tag{2}$$

[3]We consider the reference case when players select based on the viewcount only for the sake of explanation. We will extend the model to other interesting cases in next sections.

[4]A traditional application of Wardrop equilibria is road traffic, where users tend to settle to routes minimizing their delay: the effect of a route change of an individual driver belonging to a flow is negligible system-wide to the utilities of other users.

[5]In a single source diffusion model, for instance, $X = N(1 - \exp(-\lambda t)) = N\lambda t + o(t)$

If $\alpha \leq \beta$, then $X(t,\theta) = X_{ps}(t,\theta)$ for $0 \leq t \leq t_\alpha$ and $X(t,\theta) = X_{ps}(t,\theta) + X_{pu}(t,\theta)$ for $t_\alpha \leq t \leq t_\beta$. In this case,

$$t_\alpha = \frac{\alpha}{\lambda_{pu}(\theta)}, \quad t_\beta = \frac{\beta - \alpha}{\lambda_{ps}(\theta) + \lambda_{pu}} + t_\alpha$$

and in turn

$$\begin{aligned} U(\alpha,\beta) \quad &= \tau(\pi_G - \pi_B) - \alpha\Big(\frac{\pi_G}{\lambda_{ps}(G)} - \frac{\pi_B}{\lambda_{ps}(B)}\Big) \\ &-(\beta-\alpha)\Big(\frac{\pi_G}{\lambda_{ps}(\theta)+\lambda_{pu}} - \frac{\pi_B}{\lambda_{ps}(\theta)+\lambda_{pu}}\Big) \end{aligned} \tag{3}$$

Now, we can distinguish the three statements in the claim:

i. $\frac{\pi_G}{\lambda_{ps}(G)} \geq \frac{\pi_B}{\lambda_{ps}(B)}$: in the first case, due to linearity, $\beta = 0$ maximizes the utility; in the second case, we observe that indeed it must hold $\pi_G \geq \pi_B$, and then

$$\pi_G \lambda_{ps}(B) - \pi_B \lambda_{ps}(G) \geq 0 \geq \lambda_{pu}(\pi_B - \pi_G)$$

so that $\frac{\pi_G}{\lambda_{ps}(G)+\lambda_{pu}} \geq \frac{\pi_B}{\lambda_{ps}(B)+\lambda_{pu}}$: in turn the utility function is maximized again if $\beta = 0$. Hence, it holds $\beta^*(\alpha) = 0$.

ii. In the first case, it is optimal to maximize β, which brings $\beta = \alpha$. In the second case, in turn it is optimal to minimize β, so that again $\beta = \alpha$. Hence, $\beta^*(\alpha) = \alpha$.

iii. In the first case, the best response is the same as in ii. In the second case, instead, it is optimal to maximize β, so that again $\beta = \beta_{\tau,B}$. However, the last term of (3) is positive and $\beta = \beta_{\tau,B}$ maximizes it. Also, by comparison with (2), indeed $\beta^*(\alpha) = \beta_{\tau,B}$ in this case. ∎

The above results provide a characterization of the possible symmetric Wardrop equilibria of the system.

Theorem 1: i. if $\frac{\pi_G}{\lambda_{ps}(G)} \geq \frac{\pi_B}{\lambda_{ps}(B)}$, then 0 is a symmetric Wardrop equilibrium

ii. if $\frac{\pi_G}{\lambda_{ps}(G)} \leq \frac{\pi_B}{\lambda_{ps}(B)}$ but $\frac{\pi_G}{\lambda_{ps}(G)+\lambda_{pu}} \geq \frac{\pi_B}{\lambda_{ps}(B)+\lambda_{pu}}$,then all $0 \leq \beta \leq \beta_{\tau,B}$ are symmetric Wardrop equilibria

iii. if $\frac{\pi_G}{\lambda_{ps}(G)} \leq \frac{\pi_B}{\lambda_{ps}(B)}$ but $\frac{\pi_G}{\lambda_{ps}(G)+\lambda_{pu}} < \frac{\pi_B}{\lambda_{ps}(B)+\lambda_{pu}}$, then $\beta_{\tau,B}$ is a symmetric Wardrop equilibrium

It is possible to interpret the above result as follows: $\frac{\pi_G}{\lambda_{ps}(G)}$ represents the time pace at which push users are believed to access a good content. Similarly $\frac{\pi_B}{\lambda_{ps}(B)}$ represents the time pace at which push users are believed to access a bad content. Thus, condition i. suggests that it is always convenient to anticipate the access to the content. In case ii., the situation is dictated by the uptake of pull users, because they increase the viewcount thus reinforcing the believed viewcount pace of a good content against that of a bad content. Finally, in case iii. there is no incentive in accessing the content.

B. *Exponential case: fixed time horizon*

Let us consider the content dissemination process operated by a content provider using a finite set of potential target users. After the content is posted by the provider directly to users, it will be transmitted to more and more users by using some preferential channels. In this case, we need to model the push dynamics accounting for the size N of the pool of push users, i.e., we assume that the content provider disseminates the content according to

$$\dot{X}_{ps}(t,\theta) = \lambda_{ps}(\theta)(N - X_{ps}(t,\theta)),$$

so that

$$X_{ps}(t,\theta) = N(1 - e^{-\lambda_{ps}(\theta)t}) \text{ for } t \geq 0 \tag{4}$$

We reported in Fig. 3 the shape of the utility function under the exponential case for a fixed time horizon. As it can be observed in case a), for smaller values of α, i.e, $\alpha = 400$ a low value of the belief π_G causes the access to be delayed till time τ, whereas for increasing values of π_G we observe first a local maximum at α ($\pi_g = 0.75$), and finally the strategy $\beta = 0$ takes over corresponding to very large values of π_G. Indeed, such a behavior of the utility function resembles – for a fixed N – what we observed in the linear case. However, at a closer look, namely in Fig. 3c) we understand that the situation is more elaborate: in particular, we know that number of push users N impacts the speed at which the viewcount increases. As such, a small N does not permit to pass the threshold α, whereas a very large one incentivizes early access: recall that $\beta_{\max} := \beta_{\tau,B}$ means access at time $t = 0$. In between, the presence of a maximum predicts, as in the linear case, the existence of best responses that lie in the interior of $[0, \beta_{\max}]$. This intuitive numerical insight is confirmed by the theoretical results that we detail in the following.

We distinguish two cases, namely $\alpha < \beta$ and $\beta \leq \alpha$.

If $\beta \leq \alpha$, we have

$$t_\beta(\theta) = -\frac{1}{\lambda_{ps}(\theta)} \log\Big(1 - \frac{\beta}{N}\Big), \quad t_\alpha(\theta) = -\frac{1}{\lambda_{ps}(\theta)} \log\Big(1 - \frac{\alpha}{N}\Big)$$

Hence the utility becomes

$$U(\alpha,\beta) = (\pi_G - \pi_B)\tau + \log\Big(1 - \frac{\beta}{N}\Big)\Big(\frac{\pi_G}{\lambda_{ps}(G)} - \frac{\pi_B}{\lambda_{ps}(B)}\Big)$$

Let $\beta_1^*(\alpha)$ (resp. $\beta_2^*(\alpha)$) be the best response to α in $[0,\alpha]$ (resp. $[\alpha, \beta_{max}]$)

Lemma 2: In the exponential case, under the assumption $\lambda_{ps}(G) > \lambda_{ps}(B)$, it holds for $\beta \leq \alpha$

- If $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ then $\beta_1^*(\alpha) = \alpha$
- If $\frac{\pi_G}{\pi_B} > \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ then $\beta_1^*(\alpha) = 0$
- If $\frac{\pi_G}{\pi_B} = \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ then for every $\beta_1^* \in [0,\alpha]$ is optimal

Proof: The proof is similar to the one developed in the linear case for $\beta \leq \alpha$. ∎

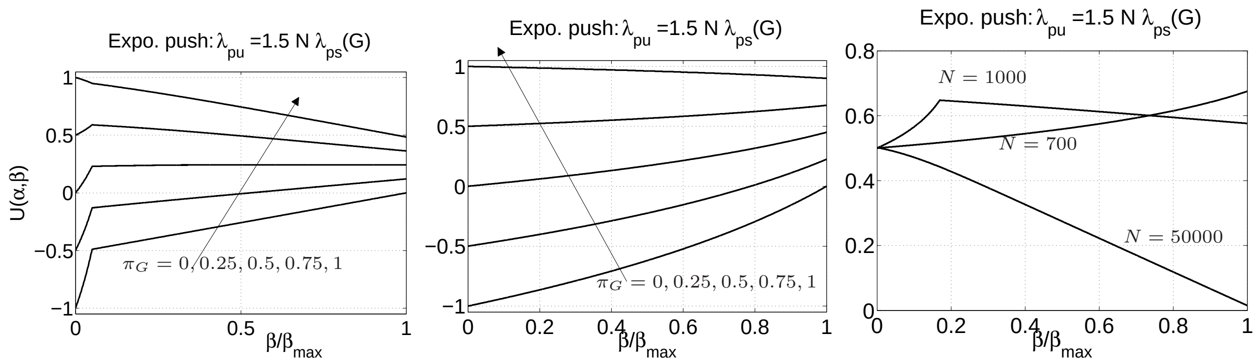
Now, we study the second case: $\alpha \leq \beta$. If $t_\alpha \leq t \leq t_\beta$,

$$X(t,\theta) = N(1 - \exp(-\lambda_{ps}(\theta)t) + \lambda_{pu}(t - t_\alpha) \tag{5}$$

for which we obtain

$$\begin{aligned} t_\beta \quad &= \lambda_{ps}(\theta)\Big(W\Big(\frac{\lambda_{ps}(\theta)}{\lambda_{pu}} N \frac{e^{\frac{\lambda_{ps}(\theta)}{\lambda_{pu}} N(1-\frac{\beta}{N})}}{\left(1 - \frac{\alpha}{N}\right)}\Big) \\ &- \log\Big(\frac{e^{\frac{\lambda_{ps}(\theta)}{\lambda_{pu}} N\left(1-\frac{\beta}{N}\right)}}{\left(1 - \frac{\alpha}{N}\right)}\Big)\Big) \end{aligned} \tag{6}$$

where $W(\cdot)$ is the Lambert function [15]. We can obtain the derivative of the above expression by letting $\xi(\beta) = \frac{e^{\zeta(\theta)(1-\frac{\beta}{N})}}{\left(1-\frac{\alpha}{N}\right)}$



(a) Case $\alpha = 400$ (b) Case $\alpha = 700$ (c) $\alpha = 700$, increasing N

Fig. 3: The utility function for $N = 1000$, for $\tau = 10$ days, $\lambda_{ps}(G) = 10^{-1}$ views/day, $\lambda_{ps}(B) = \lambda_{ps}(G)/10$. a) $\alpha = 400$ views, b) $\alpha = 700$ views. Increasing values of the belief π_G determine different shapes for the utility function. c) Increasing values of $N = 700, 1000, 50000$ for $\alpha = 700$. All graphs for $\lambda_{pu} = 1.5N\lambda_{ps}(G)$.

and $\zeta(\theta) = \frac{\lambda_{ps}(\theta)}{\lambda_{pu}}N$

$$\begin{aligned}\frac{d}{d\beta}t_\beta &= \frac{1}{\lambda_{ps}(\theta)}\frac{d}{d\beta}W(\zeta(\theta)\xi(\beta,\theta)) - \log(\xi(\beta))\\ &= \frac{1}{\lambda_{pu}}\cdot\frac{1}{1+W(\zeta(\theta)\xi(\beta,\theta))}\end{aligned}$$

After some cumbersome algebra, we derive

Lemma 3: In the exponential case, under the assumptions $\lambda_{ps}(G) > \lambda_{ps}(B)$ and $\lambda_{ps}(G)N \le \lambda_{pu}$, for $\alpha \le \beta$ it holds

- If $\pi_G \le \pi_B$ then $\beta_2^*(\alpha) = \beta_{\tau,B}$
- If $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \ge \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau,B}]$ then $\beta_2^*(\alpha) = \alpha$
- If $\frac{1+W(\zeta(G)\xi(\beta_\tau,G))}{1+W(\zeta(B)\xi(\beta_tau,B))} \le \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_\tau(B)]$ then $\beta_2^*(\alpha) = \beta_{\tau,B}$
- otherwise $\beta_2^*(\alpha)$ is the solution of the following equation

$$\frac{1+W(\zeta(G)\xi(\beta_2^*(\alpha),G))}{1+W(\zeta(B)\xi(\beta_2^*(\alpha),B))} = \frac{\pi_G}{\pi_B}$$

Proof: The derivative of the utility function U is

$$U'(\alpha,\beta) = \frac{1}{\lambda_{pu}}\Big(\frac{\pi_B}{W(\zeta(G)\xi(\beta,B))} - \frac{\pi_B}{W(\zeta(G)\xi(\beta,G))}\Big) \tag{7}$$

Since $\xi(\beta,G) > \xi(\beta,B)$ and $\zeta(G) > \zeta(B)$ then it is easy to check under condition $\pi_G \le \pi_B$ that $U'(\alpha,\beta) > 0$. Hence the utility function attains a unique maximum at $\beta_{\tau,B}$.

In order to complete the proof, it is sufficient to show that the function U is either non-increasing, or there is some $\bar\beta$ such that U is non-decreasing for $\beta < \bar\beta$ and non-increasing for $\beta > \bar\beta$.

Assume that there exists a $\bar\beta$ such that $U'(\alpha,\bar\beta) \le 0$. From (7), it is sufficient to show that

$$U'(\alpha,\beta) \le 0 \quad \text{for all } \beta > \bar\beta$$

We can show the above propriety by letting $\bar W(\beta) = \frac{1+W(\zeta(G)\xi(\beta,G))}{1+W(\zeta(B)\xi(\beta,B))}$ and it turns out that

$$\begin{aligned}\frac{\partial\bar W(\beta)}{\partial\beta} &= \frac{1}{(1+W(\zeta(B)\xi(\beta,B)))^2}\\ &\Big(\frac{\zeta(B)W(\zeta(B)\xi(\beta,B))(1+W(\zeta(G)\xi(\beta,G)))}{1+W(\zeta(B)\xi(\beta,B))}\\ &-\frac{\zeta(G)W(\zeta(G)\xi(\beta,G))(1+W(\zeta(B)\xi(\beta,B)))}{1+W(\zeta(G)\xi(\beta,G))}\Big)\end{aligned}$$

To show $\frac{\partial\bar W(\beta)}{\partial\beta} \le 0$, we impose the inequality

$$\frac{\zeta(B)W(\zeta(B)\xi(\beta,B))}{(1+W(\zeta(B)\xi(\beta,B)))^2} \le \frac{\zeta(G)W(\zeta(G)\xi(\beta,B))}{(1+W(\zeta(G)\xi(\beta,G)))^2} \tag{8}$$

We can obtain the above inequality under assumption $\lambda_{ps}(G)N \le \lambda_{pu}$ by letting

$$f(y) = \frac{yW(y\frac{e^{y(1-\frac{\beta}{N})}}{(1-\frac{\alpha}{N})})}{(1+W(y\frac{e^{y(1-\frac{\beta}{N})}}{1-\frac{\alpha}{N}}))^2}$$

Hence the derivative of f can be expressed as

$$\frac{\partial f}{\partial y} = w(\bar y)\frac{w^2(\bar y) + w(\bar y)(1-y(1-\frac{\beta}{N})) + 2 + y(1-\frac{\beta}{N})}{(1+w(\bar y)^2} \tag{9}$$

where $\bar y = y\frac{e^{y(1-\frac{\beta}{N})}}{(1-\frac{\alpha}{N})}$. In fact it can be showed that $\dot f$ is positive for $y(1-\frac{\beta}{N}) \le 1$ i.e., $\lambda_{ps}(G)N \le \lambda_{pu}$. ■

Overall, the above cases are summarized in the following theorem

Theorem 2: Let $\lambda_{ps}(G) > \lambda_{ps}(B)$ and $\lambda_{ps}(G)N \le \lambda_{pu}$, then in the exponential case

i) If $\pi_G \le \pi_B$ then $\beta_{\tau,B}$ is a symmetric Wardrop equilibrium
ii) If $\pi_G > \pi_B$ then the following cases hold
  a) If $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and $\frac{1+W(\zeta(G)\xi(\alpha,G))}{W(\zeta(B)\xi(\alpha,B))} \ge \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau,B}]$ then all $0 < \beta \le \beta_{\tau,B}$ are symmetric Wardrop equilibria

b) If $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and $\frac{1+W(\zeta(G)\xi(\beta_\tau,G))}{1+W(\zeta(B)\xi(\beta_\tau,B))} \leq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau,B}]$ then $\beta_{\tau,B}$ is a symmetric Wardrop equilibrium

c) If $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and there exists a $\bar{\beta}$ is the solution of the following equation

$$\frac{1+W(\zeta(G)\xi(\bar{\beta},G))}{1+W(\zeta(B)\xi(\bar{\beta},B))} = \frac{\pi_G}{\pi_B}$$

then $\bar{\beta}$ is a symmetric Wardrop equilibrium

iii) If $\frac{\pi_G}{\pi_B} > \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$, then the following cases hold

a) if $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \geq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau,B}]$ then 0 is a symmetric Wardrop equilibrium

b) if $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \leq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau,B}]$, then there exists a symmetric Wardrop equilibrium which is given by

$$\begin{cases} 0 & \text{if } \tau\pi_B < \pi_G t_{\beta_{\tau,B}}(G) \\ \beta_{\tau,B} & \text{if } \tau\pi_B > \pi_G t_{\beta_{\tau,B}}(G) \\ \beta^* \in \{0, \beta_{\tau,B}\} & \text{if } \tau\pi_B = \pi_G t_{\beta_{\tau,B}}(G) \end{cases} \tag{10}$$

Theorem. 2 displays a structure of the best response that is similar to the result obtained for the linear case, but we should highlight some differences. First, the additional request $\lambda_{ps}(G)N \leq \lambda_{pu}$ is excluding the case when the effect of the pull mechanism is negligible compared to push mechanism. This means that we are restricting to the case when the aggregated maximum rate at which the viewcount can increase due to the push mechanism is smaller than the increase that is generated once the viewcount is above threshold for pull users. Indeed, this is the interesting case when the content provider's aim is to attract a large basin of pull users using a target limited audience of push users.

Second, we observe that the term $\frac{\pi_\theta}{\lambda_{ps}(\theta)+\lambda_{pu}}$ that was present in the linear case is now replaced by a term involving the Lambert function $W(\cdot)$ [15]: this is due to the combined effect of the exponential growth and the linear growth above the threshold, accounting for the saturation of the basin of push users. In the case when N is very large or λ_{ps} is very small, the term collapses to the condition expressed in the linear case.

IV. Variable time horizon

In this section, we are interested in the case where the time horizon during which the content is accessed by pull users is not fixed. But, it is determined by the popularity of the content and by the quality perceived by users. In particular, when the popularity of a content is subject to saturation, we can model a vanishing $\dot{X}$ to encode the condition when a content which is present online for a long time becomes stale. Conversely, fresh uptaking contents will experience large values of $\dot{X}$ and will be preferred. This case fits well specific types of contents such as news or pop songs, for which the *trend* of the viewcount increase may be the main trigger for the users' interest in some content. Pull users still adopt a threshold strategy and browse the content if

$$\dot{X}(t,\theta) \geq \gamma_{th} \tag{11}$$

Let us consider the exponential push case introduced in the previous section. Condition (11) determines a variable horizon to access content θ:

$$\tau(\alpha,\theta) = \dot{X}^{-1}(\gamma_{th})$$

Because the time horizon $\tau = \infty$ for $\gamma_{th} \leq \lambda_{pu}$, we restrict our analysis to the case when $\gamma_{th} > \lambda_{pu}$.

Again, we are interested to compute the utility function for a tagged user given a certain common threshold strategy α played by other users; the objective is to compute the best response β for the tagged user as done before. Let $X_{th}(\theta) = N - \frac{\gamma_{th}}{\lambda_{ps}(\theta)}$, $\tau_0(\theta) = \frac{1}{\lambda_{ps}(\theta)} \log\left(\frac{\lambda_{ps}(\theta)N}{\gamma_{th}}\right)$ and $\tau_1(\theta) = \frac{1}{\lambda_{ps}(\theta)} \log\left(\frac{\lambda_{ps}(\theta)N}{\gamma_{th}-\lambda_{pu}}\right)$.

Observe that the interval of time when pull users will access the content becomes now $[\tau_0(\theta)\,\tau_1(\theta)]$: the duration of such interval corresponds to the useful lifetime of the content as dictated by the interest of the users based on (11) and by the content type.

We distinguish again two intervals, namely $0 \leq \beta \leq \alpha$ and $\alpha \leq \beta \leq \tau$, and denote β_1^* and β_2^* the best response in those intervals, respectively. However, we need to account also for (11) and to detail the utility accordingly.

It follows that if $\beta \geq \alpha$, then

$$U(\alpha,\beta) = \pi_G\Big(\tau_1(G) - t_\beta(G)\Big)^+ - \pi_B\Big(\tau_1(B) - t_\beta(B)\Big)^+$$

If $\alpha > X_{th}(G)$ and $\beta \leq \alpha$

$$\begin{aligned} U(\alpha,\beta) \;&=\; \pi_G\Big(\tau_0(G) - t_\beta(G)\Big)^+ + \pi_G\Big(\tau_1(G) - t_\alpha(G)\Big)^+ \\ &\quad -\pi_B\Big(\tau_0(B) - t_\beta(B)\Big)^+ - \pi_B\Big(\tau_1(B) - t_\alpha(B)\Big)^+ \end{aligned}$$

If $X_{th}(B) \leq \alpha \leq X_{th}(G)$ and $\beta \leq \alpha$, then

$$\begin{aligned} U(\alpha,\beta) \;&=\; \pi_G\Big(\tau_1(G) - t_\beta(G)\Big) \\ &\quad -\pi_B\Big(\tau_0(B) - t_\beta(B)\Big)^+ - \pi_B\Big(\tau_1(B) - t_\alpha(B)\Big)^+ \end{aligned}$$

If $X_{th}(B) \geq \alpha$ and $\beta \leq \alpha$, then

$$U(\alpha,\beta) \;=\; \pi_G\Big(\tau_1(G) - t_\beta(G)\Big) - \pi_B\Big(\tau_1(B) - t_\alpha(B)\Big)$$

With a similar analysis as that employed in the proof of Thm.3, we can write:

Theorem 3: In the exponential case, under the assumptions $\lambda_{ps}(G) > \lambda_{ps}(B)$ and $\lambda_{ps}(G)N \leq \lambda_{pu}$, it holds

- If $\pi_G \leq \pi_B$ then β is a symmetric Wardrop equilibrium where $\beta = \beta_\tau(B)$ is solution of $t_\beta(B) = \tau$
- If $\pi_G > \pi_B$, $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \geq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau_{0,B}}]$ then all values in the interval $[0, \tilde{\beta}]$ are symmetric Wardrop equilibria.
- If $\pi_G > \pi_B$, $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and $\frac{1+W(\zeta(G)\xi(\beta_\tau,G))}{1+W(\zeta(B)\xi(\beta_\tau,B))} \leq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau_{0,B}}(B)]$ then $\tilde{\beta}$ is a symmetric Wardrop equilibrium.

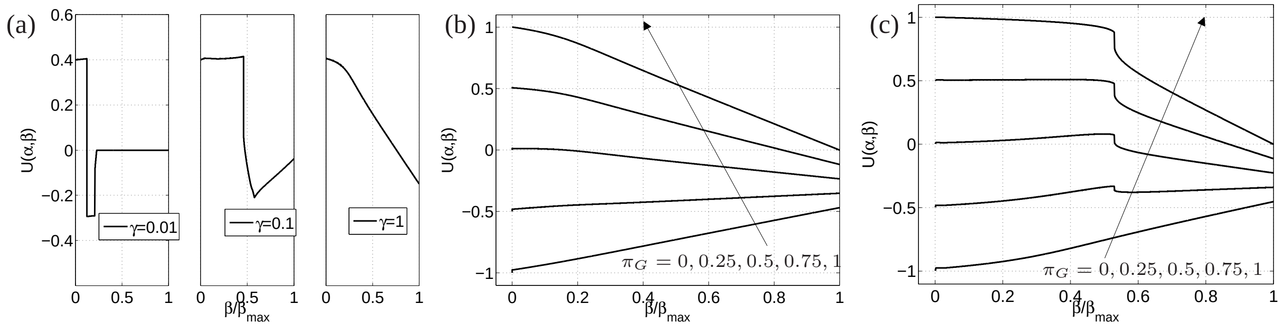


Fig. 4: The utility function for $N = 1000$, for $\tau = 10$ days, $\lambda_{ps}(G) = 10^{-1}$ views/day, $\lambda_{ps}(B) = \lambda_{ps}(G)/10$. a) Detail of the discontinuities of $U(\alpha, \beta)$ for $\gamma = 0.01, 0.1, 1$, where $\alpha = 0.18$ b) Extremal type of best response for $\alpha = 0.029$, $\gamma = 1.5$ and under increasing values of the belief π_G. c) Same as b) but for $\gamma = 0.3$. Discontinuity in α corresponds to local maxima for $\pi_G = 0.25, 0.50$.

- If $\pi_G > \pi_B$, $\frac{\pi_G}{\pi_B} < \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and there exists β_s solution of the following equation
$$\frac{1 + W(\zeta(G)\xi(\beta_s, G))}{1 + W(\zeta(B)\xi(\beta_s, B))} = \frac{\pi_G}{\pi_B}$$
then β_s is a symmetric Wardrop equilibria.
- If $\frac{\pi_G}{\pi_B} > \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and and $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \geq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau_{0,B}}]$ then 0 is a symmetric Wardrop equilibrium.
- If $\frac{\pi_G}{\pi_B} > \frac{\lambda_{ps}(G)}{\lambda_{ps}(B)}$ and and $\frac{1+W(\zeta(G)\xi(\alpha,G))}{1+W(\zeta(B)\xi(\alpha,B))} \leq \frac{\pi_G}{\pi_B}$ for all $\beta \in [\alpha, \beta_{\tau_{0,B}}]$ then there exists a symmetric Wardrop equilibrium which is given by
$$\begin{cases} 0 & \text{if } \tau\pi_B < \pi_G t_{\beta_{\tau_{0,B}}} \\ \beta_\tau(B) & \text{if } \tau\pi_B > \pi_G t_{\beta_{\tau_{0,B}}} \\ \beta^* \in \{0, \beta_{\tau_0}(B)\} & \text{if } \tau\pi_B = \pi_G t_{\beta_{\tau_{0,B}}} \end{cases} \tag{12}$$

The overall result in Thm.3 shows a structure that is close to that obtained in Thm. 2. We can conclude that the presence of a selective preference expressed in terms of the viewcount trend does not affect the structure of the Wardrop equilibria. In fact, they are of the kind determined before in the case of a fixed length interval: either extremal ones or a continuum of such restpoints. It is interesting to notice that this is following irrespective of the fact that the utility function is linear as a function of the "viewing time", i.e., the time that is useful for the viewers, but, pull users' preferences depend on a non-linear function of the threshold type.

V. Combined effect of Trend and Viewcount

In general, contents that are present online since a long time display different popularity than contents which last only a short time [7]. As we noticed in the previous section, when popularity saturation occurs, $\dot{X}$ vanishes for large t. If users choose among contents with different trend and different viewcount, they would naturally choose a content with large viewcount and large increasing trend. To this respect $y(t) = \dot{X}(t)X(t)$ encodes the condition when the pull user still values the viewcount, but, she favors a large increasing trend given two contents with the same viewcount.

Symmetric equilibria can be determined when in the system all users adopt a strategy
$$\alpha := y(t_\alpha), \quad 0 \leq t_\alpha \leq \tau$$
and again we determine the best response for a user deviating using $\beta := y(t_\beta)$ as a reply, where $0 \leq t_\beta \leq \tau$.

It is easy to see that in the linear case, the model developed in the previous section applies as long as one replaces the dynamics with the one below
$$X_{ps}(t, \theta) = \lambda_{ps}^2(\theta)t + \lambda_{ps}(\theta), \quad X_{pu} = \lambda_{pu}^2(t - t_\alpha) \cdot \mathbb{1}(t - t_\alpha)$$
so that all the results can be specialized accordingly replacing λ_{ps} and λ_{pu} with λ_{ps}^2 and λ_{pu}^2 wherever they appear. The intuition is that when the regime of content diffusion is linear, i.e., when a large number of push users exists, the trend of popularity has the only effect to reinforce the inequality $\lambda_{ps}(B) \neq \lambda_{ps}(G)$. We then move to a more interesting case.

A. Exponential push case

In the exponential case, the dynamics again is the same captured by (4), (5). We can specialize the analysis to the two cases as done before. If $\alpha \geq \beta$, $y(t_\beta) = \beta$ implies that
$$\beta = \lambda_{ps}(\theta)N^2(1 - e^{-\lambda_{ps}(\theta)t_\beta})e^{-\lambda_{ps}(\theta)t_\beta}$$
where the solution is such that $t_\beta = -\frac{1}{\lambda(\theta)} f(\beta, \theta)$ where we let $f(\beta, \theta) := \log\left(\frac{1}{2}\left(1 + \sqrt{1 - \frac{4\beta}{\lambda_{ps}(\theta)N^2}}\right)\right)$.
$$U(\alpha, \beta) = (\pi_G - \pi_B)\tau + \left(\frac{\pi_G}{\lambda_{ps}(G)} f(\beta, G) - \frac{\pi_B}{\lambda_{ps}(B)} f(\beta, B)\right)$$
After observing that $f(0, \theta) = 0$ and $f(\beta, G) \leq f(\beta, B) \leq 0$, again we obtain two extremal cases: when $\frac{\pi_G}{\lambda_{ps}(G)} \geq \frac{\pi_B}{\lambda_{ps}(B)}$ then $U(\alpha, \beta) - (\pi_G - \pi_B)\tau \leq 0$ so that $\beta = 0$ maximizes the utility. In the opposite case, namely, $\frac{\pi_G}{\lambda_{ps}(G)} \leq \frac{\pi_B}{\lambda_{ps}(B)}$, $U(\alpha, \beta) - (\pi_G - \pi_B)\tau \geq 0$, so that $\beta = \alpha$ does.

If $\alpha \leq \beta$, the condition for
$$\beta = X(t)\left(\lambda_{ps}(\theta)Ne^{-\lambda_{ps}(\theta)} + \lambda_{pu}\right)$$
gives: $$t_\beta = t_\alpha(\theta) - \frac{N}{\lambda_{pu}}\left[1 - \frac{W(f(\beta, B)\xi(B)e^{-\xi(B)})}{\xi(B)e^{-2\xi(B)}}\right]$$

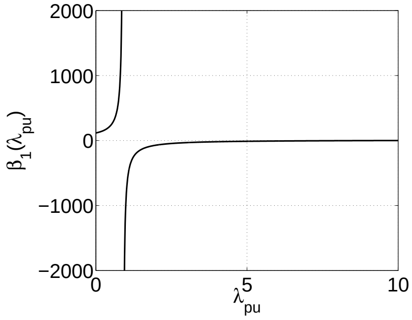


Fig. 5: The shape of function $\beta_1(\lambda_{pu})$ for increasing values of λ_{pu}: the vertical asymptote corresponds to the value λ_{pu}^s.

where we used the definition of $We^W = x$ and we stressed the dependence of t_α on θ. It is important to notice that in this case, t_β is not continuous, so that in correspondence of $t_\alpha(G)$ and $t_\alpha(B)$ the utility function has possibly two discontinuities. We reported in Fig. 4(a) the shape of the utility function for increasing values of $\gamma = 0.01, 0.1, 1$ $\lambda_{pu} = \gamma N \lambda_{ps}(G)$. For larger values of γ the effect of discontinuities becomes negligible with respect to the shape of the utility function (indeed we are looking for the best response, i.e., the maximum of $U(\alpha,\beta)$).

In particular, we observe in Fig. 4(b) that for the choice of parameters there, i.e., $\gamma = 1.5$, the shape of the utility function leads again to the customary extremal type of best response that we observed in the linear case. That is, access at time $t = 0$, i.e., $\beta = \beta_{\max}$ for large π_G and access at time $t_\beta = \tau$, i.e., $\beta = 0$ for smaller values of π_G. However, for $\gamma = 0.3$, see Fig. 4(c), we find Wardrop equilibria ($\beta^*(\alpha) = \alpha$) in the interior of $[0, \beta_{\max}]$. Further numerical exploration confirmed that the equilibria form an interval. Thus, again, we find that there exist conditions (in this case, smaller λ_{pu}) when the system has a continuum of equilibria as in previous cases.

VI. Users with side information

In the previous section we have considered the product of the trend and magnitude of the viewcount as a metric: as seen there, the structure of the equilibria that we can expect resembles closely what we found in the previous cases: either extremal Wardrop equilibria or a continuum of restpoints. We want to describe the case when potential viewers may be provided additional information on the upcoming popularity of a certain content, e.g., relying on some predictors or some apriori information they have. They judge whether to access or not a given content based on the product of the popularity X and the popularity trend $\dot{X}$. But, they only know how such metric is going to accumulate over time, i.e., the metric for a user that approaches the content at time t is

$$y(t) = \int_t^\tau X(u)\dot{X}(u)du = \frac{1}{2}(X^2(\tau) - X^2(t))$$

This metric can be used as a simple benchmark case: it contains information on the future dynamics of $X(\theta)$, and it is defined by the current and the final values of the viewcount. However, the amount of such information in general is not sufficient at time t to state the type of the content. Of course, more sophisticated metrics are possible. Nevertheless, the one at hand will do for the purpose of showing that by making the potential viewers of a content aware of some side information, the system may experience a deep change in the structure of the equilibria.

Let all users adopt strategy

$$\alpha := y(t_\alpha), \quad 0 \le t_\alpha \le \tau$$

and in the same way as done before we want to determine the best response for a user adopting $\beta := y(t_\beta)$ as a reply, where $0 \le t_\beta \le \tau$.

In the case $\beta \ge \alpha$, we recall that the dynamics is

$$X(t,\theta) = \alpha + \lambda(\theta)(t - t_\alpha)$$

where $\lambda(\theta) := (\lambda_{pu} + \lambda_{ps}(\theta))$ for the sake of notation, so that

$$\alpha + \lambda(\theta)(t_\beta - t_\alpha) = \sqrt{X^2(\tau,\theta) - 2\beta}$$

which solves for $t_\beta = \dfrac{1}{\lambda(\theta)}\Big(\alpha\dfrac{\lambda_{pu}}{\lambda_{ps}(\theta)} + \sqrt{X^2(\tau,\theta) - 2\beta}\Big)$. The corresponding expression for the utility is $U(\alpha,\beta) =$

$$U_0(\alpha,\beta) - \left[\frac{\pi_G\sqrt{X^2(\tau,G) - 2\beta}}{\lambda(G)} - \frac{\pi_B\sqrt{(X^2(\tau,B) - 2\beta)}}{\lambda(B)}\right]$$

where the term $U_0(\alpha,\beta) = (\pi_G - \pi_B)\tau - \alpha\lambda_{pu}\Big(\frac{\pi_G}{\lambda_{ps}(G)\lambda(G)} - \frac{\pi_B}{\lambda_{ps}(B)\lambda(B)}\Big)$ and it turns out that

$$\frac{dU(\alpha,\beta)}{d\beta} = \frac{\pi_G}{\lambda(G)(X^2(\tau,G) - 2\beta)^{\frac{1}{2}}} - \frac{\pi_B}{\lambda(B)(X^2(\tau,B) - 2\beta)^{\frac{1}{2}}}$$

which is decreasing with $\beta \in [-\infty, \beta_{\tau,B}]$, where $\beta_{\tau,B} := \frac{1}{2}X(\tau,B)$ as follows by comparing the ratio of the two positive terms appearing in the expression above under the assumption $X(\tau,G) \ge X(\tau,B)$). When $\frac{\pi_G}{\lambda(G)} \ne \frac{\pi_B}{\lambda(B)}$ the $U(\cdot,\beta)$ over $\mathbb{R}$ attains a unique maximum at

$$\beta_1 = \frac{1}{2}\frac{-X^2(\tau,G)\big(\frac{\pi_B}{\lambda(B)}\big)^2 + X^2(\tau,B)\big(\frac{\pi_G}{\lambda(G)}\big)^2}{\big(\frac{\pi_G}{\lambda(G)}\big)^2 - \big(\frac{\pi_B}{\lambda(B)}\big)^2}$$

so that there exists also one maximum of $U(\alpha,\beta)$ in $[t_\alpha, \tau]$.

We can distinguish three cases based on the fact that

1) $\beta_1 \le \alpha$: the best response in this case is $\beta^*(\alpha) = \alpha$
2) $\alpha < \beta_1 < \beta_{\tau,B}$: the best response is $\beta^*(\alpha) = \beta_1$
3) $\beta_1 \ge \beta_{\tau,B}$: the best response in this case is $\beta^*(\alpha) = \beta_{\tau,B}$.

Finally, we notice that when $\frac{\pi_G}{\lambda(G)} = \frac{\pi_B}{\lambda(B)}$, case 1) applies.

In the case $\beta < \alpha$, we can derive a similar analysis starting from the dynamics $X(t,\theta) = \lambda_{ps}(\theta)t$, so that

$$\beta = y(t_\beta) = \frac{1}{2}\Big(X^2(\tau,\theta) - \lambda_{ps}^2(\theta)t_\beta^2\Big)$$

so that $t_\beta = \sqrt{X^2(\tau,\theta) - 2\beta}$, and

$$\begin{aligned} U(\alpha,\beta) &= (\pi_G - \pi_B)\tau \\ &- \left[\frac{\pi_G\sqrt{X^2(\tau,G) - 2\beta}}{\lambda_{ps}(G)} - \frac{\pi_B\sqrt{(X^2(\tau,B) - 2\beta)}}{\lambda_{ps}(B)}\right] \end{aligned}$$

In turn, we can recognize the same structure for the best response as in the previous case, where the maximum of $U(\cdot, \beta)$ (when $\frac{\pi_G}{\lambda_{ps}(G)} \neq \frac{\pi_B}{\lambda_{ps}(B)}$), over $\mathbb{R}$ is attained at

$$\beta_2 = \frac{1}{2}\frac{-X^2(\tau, G)\big(\frac{\pi_B}{\lambda_{ps}(B)}\big)^2 + X^2(\tau, B)\big(\frac{\pi_G}{\lambda_{ps}(G)}\big)^2}{\big(\frac{\pi_G}{\lambda_{ps}(G)}\big)^2 - \big(\frac{\pi_B}{\lambda_{ps}(B)}\big)^2}$$

and the three cases write

1) $\beta_2 \leq 0$: the best response in this case is $\beta^*(\alpha) = 0$.
2) $0 < \beta_2 < \alpha$: the best response is $\beta^*(\alpha) = \beta_2$.
3) $\beta_2 \geq \alpha$: the best response is $\beta^*(\alpha) = \alpha$.

Again, when $\frac{\pi_G}{\lambda_{ps}(G)} = \frac{\pi_B}{\lambda_{ps}(B)}$, case 1) applies.

Now, to complete our analysis, we need to determine the best response between the two cases: we need to detail the relation between β_1 and β_2. To so do we can rewrite for the sake of convenience

$$\beta_1(x) = \frac{1}{2}\frac{\pi_G^2 x^2 X^2(\tau, B) - \pi_B^2(L + x)^2 X^2(\tau, G)}{\pi_B^2 x^2 - \pi_G^2(L + x)^2}$$

where $L = \lambda_{ps}(G) - \lambda_{ps}(B)$ and $x = \lambda_{ps}(G) + \lambda_{pu}$. It can be easily showed that

$$\frac{d}{dx}\beta_1(x) = \pi_G^2 \pi_B^2 \frac{2Lx(X^2(\tau, G) - X^2(\tau, B))(x - L)}{(\pi_B^2 x^2 - \pi_G(L + x)^2)^2}$$

which brings $\frac{d}{dx}\beta_1(x) > 0$ for $x \geq 0$, with a singularity in

$$\lambda_{pu}^s = \frac{\pi_B}{\pi_G - \pi_B}(\lambda_{ps}(G) - \lambda_{ps}(B)) - \lambda_{ps}(B)$$

The typical shape of β_1 is reported in Fig. 5. We observe that $\beta_1(\lambda_{pu} = 0) = \beta_2$. The asymptotic value for $\lambda_{pu} = \infty$ is

$$\beta_1(\infty) = \frac{1}{2}\frac{\pi_G^2 X^2(\tau, B) - \pi_B^2 X^2(\tau, G)}{\pi_G^2 - \pi_B^2}$$

It can be verified that $\beta_1(\lambda_{pu})$ is injective. Hence, the above analysis let us state: $\beta_1(\infty) \leq \beta_1(0) = \beta_2$, which in turn leads to the following

Lemma 4: For $0 \leq \lambda_{pu} < \lambda_{pu}^s$, it holds $\beta_1 \geq \beta_2$, and for $\lambda_{pu} > \lambda_{pu}^s$ it holds $\beta_1 < \beta_2$.

Now we can combine the conditions above to derive:

Theorem 4: Let $I = [0, \beta_{\tau,B}]$

i. If $\lambda_{pu} > \lambda_{pu}^s$, then

$$W_s = [\beta_1, \beta_2] \cap I$$

is the set of symmetric Wardrop equilibria for the system.

ii. If $\lambda_{pu} < \lambda_{pu}^s$ then $W_s \subseteq \{0, \beta_{\tau,B}\}$.

Proof: Case i. follows immediately observing that for $\beta_1 \leq \alpha$ the best response is $\beta^*(\alpha) = \alpha$ and for $\beta_2 \geq \alpha$ the best response is $\beta^*(\alpha) = \alpha$: both conditions are satisfied simultaneously for $\alpha \geq 0$ if and only if $\alpha \in W_s$.

Case ii. is proved observing that the conditions for case i fail, so that only extremal cases can hold. In particular, W_s is not always the empty set: if $\beta_2 \geq 0$, then $\beta_1 \leq 0$ so that $\beta^*(0) = 0$ and the same holds in the opposite case, i.e., if $\beta_1 \geq \alpha = \beta_{\tau,B}$ then $\beta_1 \geq \alpha = \beta_{\tau,B}$ so that $\beta^*(\beta_{\tau,B}) = \beta_{\tau,B}$. ■

The result in Thm. 4 let us observe a neat phase transition effect on λ_{pu}: when the intensity of the views due to the pull mechanism is below threshold λ_{pu}^s, only extremal Wardrop equilibria are possible. Above that threshold, there can exist a continuum of equilibria where the system can settle. Let $\mu(\cdot)$ denote the standard real measure: a sufficient condition is provided in the following

Corollary 1: $\mu(W_s) > 0$ if $\lambda_{pu} > \lambda_{pu}^s$ and $\beta_2 \geq 0 > \beta_1$.

We can observe that $\pi_G < \pi_B$ implies $\beta_2 \geq 0$ and $\lambda_{pu} > 0 > \lambda_{pu}^s$, so that a stronger sufficient condition than the one just provided in turn becomes: $\pi_G < \pi_B$ and $\beta_1 \leq \beta_{\tau,B}$.

VII. Related Works

The analysis of dynamics of popularity of online contents has been subject of recent papers. The work [3] provides an analysis of the YouTube system, with comprehensive view of the characteristic of the generated traffic.

In [5] the authors address the relation between metrics used to evaluate popularity. They observed that viewcount is strongly correlated with several such metrics as number of comments, ratings, or favorites. However, all such metrics do not correlate to average rating. In this paper we confine our analysis to viewcount as the metric of interest. [7] focuses on the core problem of predicting popularity, namely, the viewcount, based on early measurements of user access. Based on YouTube videos or Digg stories measurements, the authors observe that contents increasing fast their viewcount in early stages typically become popular later on. The proposed empirical model, i.e., $\log N(t_r) = \log N(t_o) + \lambda_0(t_r, t_0)$ where $\lambda_0(t_r, t_0)$, is a random multiplicative noise and $N(t_r), N(t_0)$ is the viewcount at t_r and t_0; it resembles closely the exponential model adopted in this work.

In [4] the authors propose a model accounting for change of ranking induced by UGC online platforms. The model is meant to overcome the limitations of the preferential attachment models. Those models in fact cannot explain bursty growth of content popularity; those in turn are claimed an inherent property of the online platforms. The authors relate bursty growth spikes to the way such systems expose popular contents to users and perform re-ranking of existing contents causing positive feedback loops.

The paper [6] provides analysis of power law behavior for the rank distribution of contents; the distribution of most watched videos is found heavily skewed towards the most popular ones.

Threshold models similar to those studied in this work are described by Granovetter [11] in social science. The assumption is that individuals make binary decisions (in our framework, view or not view a content), according to some static internal threshold of others participating. A generalization based on threshold distribution is addressed in [9].

VIII. Conclusions

In this paper we characterized the access to online contents by game theoretical means by leveraging on the concept of Wardrop equilibrium. We deduced the structure of equilibria in systems where users adopt threshold type policies to select online contents. We explored several cases: the case when the plain viewcount is the metric, or the viewcount trend, or both are combined as a product metric. We explored the case of a fixed time horizon dictated by the content lifetime, and we considered a case when the time horizon is not fixed. Finally we explored the impact of side information available to users.

In all such cases we deduced the presence of a continuum of equilibria, which has potential implications in the design and control of platforms for online content access. In future work, in particular, we are exploring the dynamics associated to such sets of interior restpoints, when they exist, and comparing those with typical dynamics of online contents. However, not only equilibria are relevant: as showed in [10], threshold strategies, under specific conditions, may well lead the system to be asymptotically unstable; system trajectories may in turn consist of cycles that can move into a chaotic dynamics, essentially indistinguishable from random noise.

REFERENCES

[1] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, “I tube, you tube, everybody tubes: analyzing the world’s largest user generated content video system,” in *Proc. of ACM IMC*, San Diego, California, USA, October 24-26 2007, pp. 1–14.

[2] R. Crane and D. Sornette, “Viral, quality, and junk videos on YouTube: Separating content from noise in an information-rich environment,” in *Proc. of AAAI symposium on Social Information Processing*, Menlo Park, California, CA, March 26-28 2008.

[3] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, “YouTube traffic characterization: A view from the edge,” in *Proc. of ACM IMC*, 2007.

[4] J. Ratkiewicz, F. Menczer, S. Fortunato, A. Flammini, and A. Vespignani, “Traffic in Social Media II: Modeling Bursty popularity,” in *Proc. of IEEE SocialCom*, Minneapolis, August 20-22 2010.

[5] G. Chatzopoulou, C. Sheng, and M. Faloutsos, “A First Step Towards Understanding Popularity in YouTube,” in *Proc. of IEEE INFOCOM*, San Diego, March 15-19 2010, pp. 1 –6.

[6] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, “Analyzing the video popularity characteristics of large-scale user generated content systems,” *IEEE/ACM Transactions on Networking*, vol. 17, no. 5, pp. 1357 – 1370, 2009.

[7] G. Szabo and B. A. Huberman, “Predicting the Popularity of Online Content,” *Communications of the ACM*, vol. 53, no. 8, pp. 80–88, Aug. 2010.

[8] L. G. Debo, C. Parlour, and U. Rajan, “Signaling quality via queues,” *Manage. Sci.*, vol. 58, no. 5, pp. 876–891, May 2012.

[9] M. Rolfe, “Social networks and threshold models of collective behavior,” *Preprint, University of Chicago*, 2004.

[10] M. Granovetter and R. Soong, “Threshold models of interpersonal effects in consumer demand,” *Journal of Economic Behavior and Organization*, no. 7, pp. 83–99, 1986.

[11] M. Granovetter, “Threshold models of collective behavior,” *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.

[12] J. Wardrop, “Some theoretical aspects of road traffic research,” *Proc. Inst. Civil Eng., Part 2*, vol. 1, pp. 325–378, 1952.

[13] R. Hassin and M. Haviv, “Equilibrium threshold strategies: The case of queues with priorities,” *Oper. Res*, pp. 966–973, 1997.

[14] A. Haurie and P. Marcotte, “On the relationship between Nash-Cournot and Wardrop equilibria,” *Networks*, vol. Volume 15, no. 3, p. 295308, 1985.

[15] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, , and D. E. Knuth, “On the Lambert W Function,” *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.