Dominance or Fair Play in Social Networks? A Model of Influencer Popularity Dynamics

Franco Galante[1], Chiara Ravazzi[2], Luca Vassio[1], Michele Garetto[3], and Emilio Leonardi[1,2]

[1] Politecnico di Torino, Italy
{franco.galante, luca.vassio, emilio.leonardi}@polito.it
[2] National Research Council (CNR-IEIIT), Italy
chiara.ravazzi@cnr.it
[3] Università di Torino, Italy
michele.garetto@unito.it

Abstract. This paper presents a data-driven mean-field approach to model the popularity dynamics of users seeking public attention, i.e., influencers. We propose a novel analytical model that integrates individual activity patterns, expertise in producing viral content, exogenous events, and the platform's role in visibility enhancement, ultimately determining each influencer's success. We analytically derive sufficient conditions for system ergodicity, enabling predictions of popularity distributions. A sensitivity analysis explores various system configurations, highlighting conditions favoring either dominance or fair play among influencers. Our findings offer valuable insights into the potential evolution of social networks towards more equitable or biased influence ecosystems.

1 Introduction

Online social networks (OSNs) crucially shape digital interaction, increasingly replacing traditional media. This shift has created a fertile ground for the rise of a new class of online users: influencers. These are individuals with a large following on OSNs, who typically seek to expand it. They are also central to marketing because of their ability to drive consumer behavior [10]. Therefore, understanding how influence hierarchies form and consolidate over OSNs is of great interest. Given the growing role of platforms in shaping opinions, it is crucial to determine whether dominant users rise through their merits or skewed feedback mechanisms that concentrate users' preferences.

Despite significant interest and numerous social network analyses, a gap remains in understanding the quantitative dynamics of popularity evolution. Firstly, a universal definition of popularity in social contexts is lacking, with the literature offering varied interpretations tied to specific scientific contexts. Popularity metrics are often identified with visibility metrics [1], such as follower count or interactions [19]. However, the follower count is relatively static and its validity has been questioned [5] due to its insensitivity to interaction dynamics and rare decrease [1]. Another research direction uses a system perspective, interpreting popularity dynamically (see [4] and references therein). Among these models are epidemic [18], self-exciting (e.g., Hawkes [8]), and Bass diffusion models [2]. It is established that opinion changes result from collective interaction and mutual influence [11], but quantifying these influences is difficult [17]. Mean-field models address this by replacing pairwise interactions with an average interaction across the population [6], where individual dynamics depend on the aggregate system.

By adopting a mean-field perspective and drawing inspiration from OSN data (Facebook in this case), we propose a novel, comprehensive model that captures the evolution of real-world influencers by accounting for the following factors: (i) Collective attention, a limited resource influencers compete for [20], tends to decrease over time without new stimuli or maintenance [14]; (ii) Content creation patterns, traditionally modeled as a homogeneous Poisson process [3], are demonstrably shaped by the influencer's popularity [9]; (iii) Post success (and its impact on influencer popularity) is modeled as a random variable potentially dependent on acquired popularity [16], user characteristics (competence, experience, attractiveness), and platform feedback [13] (although undisclosed, engagement maximization algorithms are recognized to rely on past popularity and user preferences [7]); (iv) Exogenous, platform-independent events, can influence popularity [12].

Beyond introducing the model, our primary contribution is empirical evidence supporting the analytical treatment's core hypotheses. Due to the system's stochasticity, the dynamics do not deterministically converge to a single state. However, under specific conditions, we prove the system's state, described by a Markov process, is ergodic, possessing a long-run invariant distribution. This distribution, while not explicitly formulated, solves a system of partial differential equations. This result enables a probabilistic study of popularity emergence and a comparison of influencer distributions by quantifying their emergence probability and duration in privileged positions.

2 Preliminary Empirical Analysis

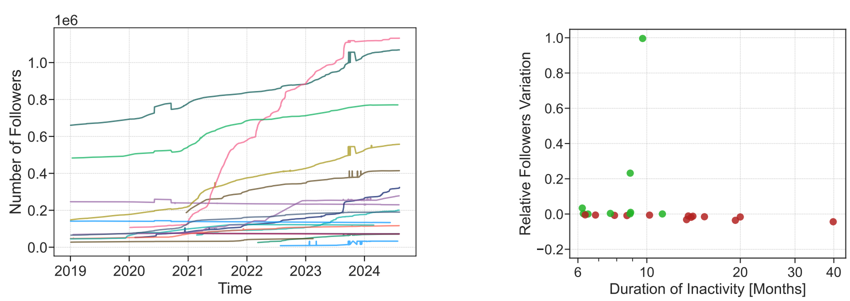
Using real data traces, we seek to characterize the popularity of influential individuals and empirically identify the key elements that contribute to the rise or fall of their prominence over time. We focus on the social network Facebook[4], where influencers publish content (i.e., posts), share content that followers, as well as other platform users, can view, like, and comment on. We restrict the analysis to influencers who post primarily on chess (as the FIDE rating provides an index of competence) and two other popular topics: cars and science. We created lists of such influencers by actively searching for them on the platform through hashtags and keywords and by consulting public lists available online.[5]

[4] A parallel analysis of Instagram data has been completed and will be presented in a forthcoming extension of this work.

[5] e.g., https://hireinfluence.com/blog/top-science-influencers/

Note that when we use the term influencer, we do not only mean physical individuals but also magazines, organizations, or companies. For each monitored influencer, we analyzed all the data related to the posts published between January 1, 2014, and May 15, 2024, using the CrowdTangle tool and its API.[6] Available data include the number of followers at the publishing time and the time-series of engagement metrics for each post. In total, we collected $1,965,805$ posts from 111 influencers.

The number of followers is a straightforward and widely adopted proxy in the literature [1] for gauging an influencer's popularity at a given time; however, it does not fully represent their popularity at that specific point. For example, we observe that the number of followers typically exhibits a purely increasing trend. Figure 1 (a) illustrates this pattern, reporting the followers' evolution on Facebook for the influencers who mainly post about chess. This phenomenon occurs because users have no real incentive to unfollow an influencer. Even when influencers stop posting for extended periods, they retain most of their followers. To illustrate this, we examine inactive influencers, that is, those who have not posted on the platform for more than 6 months. Figure 1 (b) shows the follower's variation after an inactivity period. In most cases, the follower's difference is negligible, with losses (in red) and gains (in green) occurring about equally.



(a) (b)

Fig. 1: Temporal evolution of the number of followers for *chess* influencers on Facebook (a) and a scatter plot of the relative followers' variation after a period of inactivity longer than 6 months (b), in log x-scale. Each line/dot is an influencer.

3 A Mean-Field Model of Popularity

To overcome the intrinsic limitations of the number of followers as popularity metric, we propose the following approach. We consider a finite set of influencers $\mathcal{I}$ and assign to each influencer $i \in \mathcal{I}$ a scalar variable $X^{[i]}(t) \in \mathbb{R}$, representing

[6] CrowdTangle was a public insights tool from Meta available to researchers. https://transparency.meta.com/en-gb/researchtools/other-datasets/crowdtangle/

their instantaneous *effective popularity.* $X^{[i]}(t)$ is a measure of the level of interest among platform users in the content posted by influencer $i \in \mathcal{I}$ at time $t \in \mathbb{R}$.

Based on the empirical evidence of Section 2, we assume that $X^{[i]}(t)$ is not simply proportional to the number of followers but is governed by autonomous dynamics, determined mainly by the following four factors: (i) the natural tendency of users to lose or shift their interest; (ii) the impact of competition from other influencers on the same topic, diverting attention away from influencer i; (iii) the posting activity of influencer i, aimed to secure current and attract new audiences; (iv) exogenous events affecting the online visibility of influencers, such as newsworthy events as reported by traditional media, which expose them to broader audiences beyond their established following. Formally, we assume $X^{[i]}(t)$ evolves according to the following stochastic differential equation:

$$\mathrm{d}X^{[i]}(t) = -\gamma X^{[i]}(t)dt + V_i(t)N_I^{[i]}(dt) + W_i(t)N_E^{[i]}(dt) \tag{1}$$

where each of the terms on the right-hand side corresponds to some of the previously described factors. More in detail, we first assume that $X^{[i]}(t)$ decreases at a rate γ in the absence of additional stimuli. This decaying rate is a consequence of the combined effect of: (i) lost or shifted interest, and (ii) competition. Note that we do not model the detailed interactions between influencers, but we represent the combined effect of all competitors on the same topic.

The second term represents the change in popularity due to the emission of a post. Here $N_I^{[i]}(t)$ is the counting process induced by the post-emission. To ensure generality, and as supported by empirical evidence detailed in Section 3.1, $N_I^{[i]}$ is modeled as a point process admitting a conditional intensity (also referred to as *stochastic intensity*) which may depend on the effective popularity, i.e., $\lambda_i(t) := \lambda_i\left(X^{[i]}(t)\right)$.[7] A post published at time t typically attracts audience attention, leading to an increase of influencer i's popularity (hereafter, we will informally refer to such an increase in popularity as *jump*), modeled by random variable $V_i(t)$. Note that the distribution of $V_i(t)$ depends on factors such as post quality and attractiveness, as well as the size and engagement of interacting users. Actually, the reach of a post is determined through multiple mechanisms: platform-specific algorithmic curation (such as EdgeRank) designed to optimize engagement metrics, subsequent redistribution via user sharing behaviors, and direct exposure to the influencer's established follower base. Occasionally, posts *go viral*, achieving unexpected success beyond followers. For simplicity and generality of the model, we abstract away the specific mechanisms by which posts attract attention. Instead, we limit ourselves to statistically characterizing the cumulative effect of such mechanisms, making $V_i(t)$ simply dependent on the current popularity of influencer i at time t, $X^{[i]}(t^-)$.

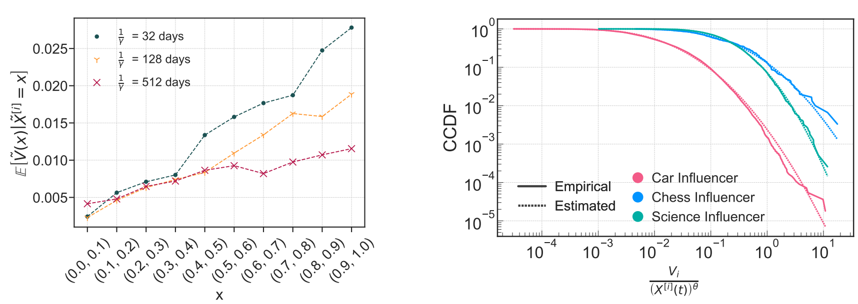
Finally (third term), external factors can affect influencer popularity. Exogenous events may cause popularity jumps, denoted by $W_i(t)$. Due to their external nature, $W_i(t)$ are reasonably independent of influencer i's behavior on the platform. The arrival process $N_E^{[i]}$ can be modeled as a Poisson process.

[7] $N_I^{[i]}$ can be made a homogeneous Poisson process by setting $\lambda_i(t) := \lambda_i$.

3.1 Statistical Characterization and Main Assumptions

We now characterize the key random variables in Eq. (1) through empirical validation on Facebook data. Then, we formalize our findings as a set of assumptions.

Post Popularity Jumps. First, we define the success of a post at time t as the number of likes (positive reactions) it receives. This number corresponds to the popularity jump V_i at time t. Then, we show its dependence on the instantaneous popularity $X^{[i]}(t^-)$. We reconstruct the popularity dynamics $X^{[i]}$ based on Eq. (1), beginning with a zero initial condition. We identify the empirical posting time sequences with the sample paths of the point process $N_I^{[i]}(dt)$, and we set $N_E^{[i]}(dt) = 0$ due to the lack of information about exogenous events. We then derive a normalized popularity $\tilde{X}^{[i]}(t)$ by dividing the popularity by its maximum value. This normalization is necessary to aggregate data from various influencers whose absolute popularity levels can differ significantly. Similarly, we use a normalized version of the empirical success $\tilde{V_i}$ of influencer i's posts. Figure 2(a) shows the average normalized jump conditioned on the instantaneous normalized popularity. We partitioned the range of normalized popularity into 10 bins and calculated the mean number of likes for posts falling within each bin. The results, presented for different values of parameter γ appearing in equation (1) ($\gamma \in \{32, 128, 512\}$ days), demonstrate a steady upward trend with increasing normalized popularity. As γ decreases (i.e., $1/\gamma$ increases), the dependency becomes smoother, likely due to the increased inertia in influencer popularity dynamics.



(a) (b)

Fig. 2: (a) Empirical conditional expectation of the normalized jump as a function of the normalized effective popularity. (b) Empirical and estimated CCDF of the ratio between the *number of likes* and $(X^{[i]})^\theta$ for three Facebook influencers.

Based on these empirical observations, we assume that $\mathbb{E}[V_i(t) \mid X^{[i]}(t^-) = x]$ is an increasing function of x. In particular, we make the following hypothesis:

Assumption 1 $\mathbb{E}[V_i(t) \mid X^{[i]}(t^-) = x] = \varepsilon + \beta_i x^\theta$, *where* ϵ *is an arbitrarily small positive constant to avoid absorbing states and* $\beta_i > 0$.

Parameter β_i is influencer-specific and reflects the influencer's competence on the topic and their ability to create captivating content. In contrast, we assume θ to be primarily determined by users' behavior and platform engagement algorithms, and thus influencer independent. It is worth mentioning that the publication of a controversial post might lead to a decrease in popularity (negative jump). However, since these events are statistically rare and hardly detectable in our dataset, we decided to neglect them.

For the conditional distribution $F_{V_i}(z \mid x)$, we make the simplest possible assumption by considering that V_i scales proportionally to $\varepsilon + \beta_i(X^{[i]}(t^-))^\theta$:

Assumption 2 *Let $\hat{V}_i$ be a positive random variable independent from $X^{[i]}(t^-)$, with unitary mean. Then $V_i = \left(\varepsilon + \beta_i(X^{[i]}(t^-))^\theta\right) \hat{V}_i$.*

It should be noticed that, according to Assumption 2, the conditional partition function is given by $F_{V_i}(z \mid x) = F_{\hat{V}_i}(\frac{z}{\varepsilon+\beta_i x^\theta})$.

Our goal is to determine the distribution of success from empirical data, again considering the number of received likes as a proxy of the post's success. To this end, we have selected a set of *candidate* distributions to be tested, including: exponential, lognormal, and power-law. A few parameters need to be first identified. We must differentiate between success-specific parameters (e.g., β_i), which reflect influencer posting behavior, and system-level parameters (i.e., $(\gamma,\ \theta)$), which capture broader user and platform dynamics. While success-specific parameters can be estimated from the data (for each chosen distribution), system-level parameters cannot be directly extracted from traces. Therefore, we have developed a Maximum Likelihood Estimation (MLE) procedure that, given both a candidate distribution and the pair (γ, θ), first finds the best success-specific parameters. Once the parameters have been obtained, we evaluate the quality of the fitting by computing the Kolmogorov distance κ between the synthetic and the empirical distribution. Recall that the Kolmogorov distance between two distributions F and G is defined as $\kappa(F, G) = \sup_{x\in\mathbb{R}} |F(x) - G(x)|$. Since the lognormal distribution results the best one (i.e., the one minimizing κ) for over 98% of the influencers under several choices of pair (γ, θ) (as better detailed below), we decided to consistently adopt the lognormal shape for distribution $\hat{V}_i$.

Recall that a lognormal distribution has two parameters. For simplicity, we neglected ε and focused on $\beta_i\hat{V}_i$, which is, as consequence of Assumption 2, a lognormal random variable with mean β_i (recall $\mathbb{E}[\hat{V}] = 1$) and coefficient of variation CV_i. The above MLE procedure can be used to estimate both parameters for each influencer. In Figure 2 (b), we report the result of the fitting for three influencers, one for each of the considered domains, comparing the empirical Complementary Cumulative Distribution Function (CCDF) with the synthetic lognormal distribution for a particular choice of (γ, θ), as discussed below.

Finally, we argued that system-level parameters (γ and θ) cannot be directly measured. However, we can determine them indirectly. First note that, given a choice of (γ, θ) and performing the MLE, it is possible to compute the Kolmogorov distance $\kappa_i(\gamma, \theta)$ between the synthetic and empirical distribution. This provides a measure of the goodness of the fit. At this point, we select the best system-level parameters as those that minimize the cumulative Kolmogorov distance (over all influencers), namely: $(\gamma^*, \theta^*) = \min_{\gamma,\theta} \sum_i \kappa_i(\gamma, \theta)$.

Figure 3 reports values of $\sum_i \kappa_i(\gamma, \theta)$ for different choices of (γ, θ) . Observe that the best fitting of the empirical data is obtained for $\theta^* = 0.7$ and $\gamma^* = 1/128$ [1/days]. These are also the parameters used in Figure 2 (b).

In Table 1, we present our estimates of β_i for five prominent influencers within the chess domain, alongside their Elo ratings (for standard chess), widely recognized as the definitive measure of chess playing strength. Notably, the correlation between β_i and Elo is not particularly strong. This suggests that technical expertise in the topic is just one of several factors contributing to an influencer's success on social media platforms.

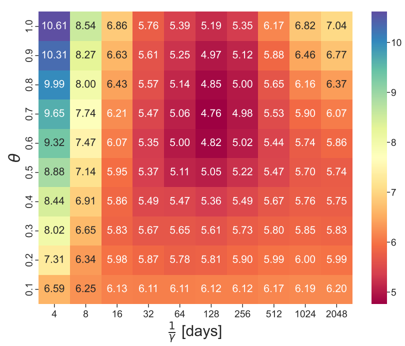


Fig. 3: Cumulative Kolmogorov distance between lognormal fit and empirical distribution.

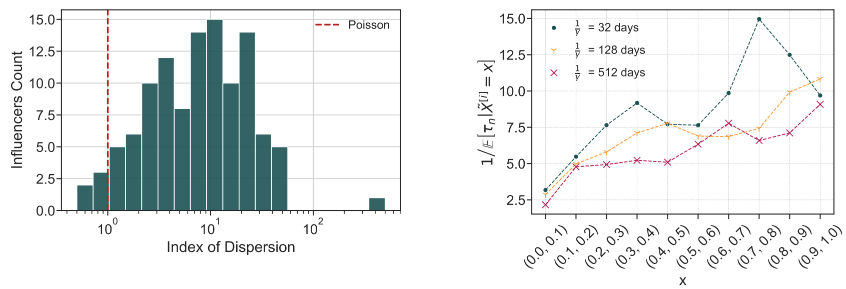
Table 1: Comparison between the influencer's estimated β_i and their FIDE Elo score.

Influencer	**Elo**	β_i
Judit Polgar	2675	0.424
Alexandra Botez	2044	0.207
Magnus Carlsen	2837	0.871
Fabiano Caruana	2776	1.194
Tania Sachdev	2396	1.953

Intensity of the Posting Process. The posting patterns of influencers represent another critical factor shaping popularity dynamics. Many studies assume the posting process to be a homogeneous Poisson process, and indeed this simple process has been found to describe well most of the activity [8]. However, data traces exhibit pronounced bursty patterns, indicating that the Poisson assumption does not always align with real-world data.

Recall that in a Poisson process, the number of events occurring within a fixed time window is distributed as a Poisson random variable. The index of dispersion of a Poisson distribution (i.e., the ratio between the variance and the mean of the distribution) is equal to 1. We computed the index of dispersion D of the number of posts in one-week windows, for each influencer. In Figure 4 (a), we report the distribution of D, whose values are typically much larger than 1, suggesting that the posting process is more complex than Poisson.

Let $\{\tau_n\}_n$ be the sequence of inter-post times, i.e., the time interval between two consecutive posts by the same influencer. Figure 4 (b) shows the empirical estimate of the reciprocal of the average next inter-post time, conditionally over the normalized (w.r.t. the maximum) instantaneous popularity $\tilde{X}^{[i]}(t)$ of all influencers. This estimation represents an approximation of the stochastic intensity $\lambda(t)|X^{[i]}(t) = x$. Even if the trend is noisier than that in Figure 2, there is a clear positive correlation between the posting rate and the effective popularity. The correlation becomes weaker as the value of γ decreases ($1/\gamma$ increases), due to the larger system inertia. Motivated by these observations, we make the following hypothesis.



(a) (b)

Fig. 4: (a) Distribution of the influencer's index of dispersion of the number of arrivals in a one-week window over the entire time horizon. (b) Instantaneous stochastic rate for different choices of γ, measured in posts/day.

Assumption 3 *For any influencer $i \in \mathcal{I}$, the conditional stochastic intensity of $N_I^{[i]}(dt)$ follows the law:* $\lambda^{[i]}(t) \mid \{X^{[i]}(t) = x\} = \lambda_{0,i} + \lambda_{1,i} x^{\phi_i}$.

This assumption maintains reasonable flexibility while capturing the observed pattern that influencers' posting rate tends to increase with growing effective popularity. However, the specific parameters characterizing this relationship vary among influencers, reflecting individual behavioral factors.

3.2 Discussion on Parameters and their Relationships

We summarize all of the model parameters in Table 2. Note that β_i can be regarded as the average strength of the content generated by influencer i, and it subsumes two different aspects: (i) the intrinsic average quality of their posts; (ii) the level of engagement of their audience, which depends on behavioral features of i, as well as, from intrinsic characteristic of their audience.

Parameter θ, instead, accounts for the effect of the platform's engagement mechanism(s) to determine the audience of individual posts (and for such a reason it is assumed independent of i). Parameters $\lambda_{0,i}$, $\lambda_{1,i}$ and ϕ_i are tightly related to behavioral traits of influencer i (like impulsivity). At last, parameter CV_i quantifies the variability in post success.

It is important to note that these parameters exhibit interdependencies. For example, we can anticipate an inverse relationship between β_i and the posting rate λ_i (determined by $\lambda_{0,i}$, $\lambda_{1,i}$ and ϕ_i): if an influencer posts too frequently, this might compromises content quality due to constrained resources and time, while potentially fatiguing their audience, leading to a reduction of post effectiveness (represented by β_i). The above natural feedback mechanism inherently discourages influencers from adopting overly aggressive posting strategies.

Furthermore, a nuanced relationship exists between the popularity decay rate γ and content publication parameters. When γ is large – whether due to natural decline of audience attention or competition from alternative content – influencers are expected to increase their posting frequency (within the previously discussed constraints) to prevent the loss of influence and visibility. Conversely, a smaller popularity decay rate reduces the urgency to post frequently, as popularity remains more stable over time, weakening the correlation between posting rate and instantaneous popularity.

Table 2: Parameters summary and reference scenario for the numerical analysis.

Symbol	**Description**	**Reference value**
β_i	ability of influencer i, $i \in \mathcal{I}$	0.9^{i-1}
γ	popularity discount rate	1/64 (days)
θ	exponent for non-linear dependence of jumps on popularity	0.6
ϵ	small constant to avoid absorption in zero	0.01
$\lambda_{0,i}$	constant term of posting rate	4 (posts/day)
$\lambda_{1,i}$	scale factor of the variable term of posting rate	0
ϕ_i	exponent of the variable term of posting rate	0
CV_i	coefficient of variation of the distribution $\hat{V}_i$	4

3.3 Theoretical Results

In this section, we carry out a complete probabilistic analysis of the effective popularity $X(t)$ as specified in Eq. (1).

First, observe that Eq. (1) defines a homogeneous Continuous Time Markov Process over a general space state. We denote with $\{X_n\}_{n\in\mathbb{N}}$ the embedded Discrete Time Markov Process (DTMP), obtained by sampling $\{X(t)\}_{t\in\mathbb{R}_+}$ at jump times, i.e., $X_n = X(T_n^-)$, where $\{T_n\}_{n\in\mathbb{N}}$ are the ordered points of $N_I + N_E$. We refer the reader to [15] for a comprehensive analysis of Markov processes over a general state space. Denoted with $\mathcal{B}(\mathbb{R}_+)$ the family of Borellian sets over $\mathbb{R}_+$, and with $P^n(x, A) = \mathbb{P}(X_n \in A \mid X_0 = x)$ for $A \in \mathcal{B}(\mathbb{R}_+)$, $n \in \mathbb{N} \setminus \{0\}$, the n-step Markov kernel of the DTMP, we have the following result.

Theorem 1. *Let Assumption 2-3 be satisfied with $\epsilon > 0$, $\lambda_0 > 0$. If $\theta + \phi < 1$, or $\theta + \phi = 1$ and $\beta(\lambda_0 + \lambda_1)$ is sufficiently small with respect to γ, then $\{X_n\}_n$ admits a unique stationary probability measure, denoted by π. Moreover, for any initial condition, as n grows large $\sup_{A\in\mathcal{B}(0,\infty)} |P^n(x, A) - \pi(A)| \to 0$.*

The detailed proof can be found in our technical report http://arxiv.org/abs/2507.03448. In short, we show μ_{Leb}-irreducibility[8], as well as strong aperiodicity for the embedded DTMC $\{X_n\}_n$, using a rather direct approach. Then, Harris recurrence can be proved using standard drift arguments, i.e., by adopting $\mathcal{L}(X_n) = X_n$ as Lyapunov function [15].

Theorem 1 guarantees that, under the appropriate conditions on system parameters, the Markov process admits a unique stationary probability measure; the process is ergodic, implying convergence to the stationary distribution regardless of the initial condition. This means the Markov chain is ergodic, and time averages of observables will converge to the expected values under the stationary distribution.

At last, with rather standard arguments, it can be shown that the ergodicity of $\{X_n\}_n$ implies the ergodicity of $\{X(t)\}$ (i.e. a unique stationary distribution $\Pi(A)$ exists and $P^t(x, A) \to \Pi(A)$ for every $x \in \mathbb{R}^+$ and $A \in \mathcal{B}(\mathbb{R}^+)$.

Denoting by $F(y,t) := \mathbb{P}(X(t) \leq y)$ the partition function of $X(t)$, $f(x,t)$ the associated probability density and μ the rate of process $N_E^{[i]}$, we have:

Theorem 2. *Under the assumptions of Theorem 1, Function $F(y,t)$ satisfies the following Partial Integro-Differential Equation*

$$\frac{\partial F(y,t)}{\partial t} = \gamma y \frac{\partial F(y,t)}{\partial y} - \int [\lambda(t)\bar{F}_V(y-x \mid x) + \mu \bar{F}_W(y-x)] \frac{\partial F(x,t)}{\partial x} \mathrm{d}x \tag{2}$$

where $\bar{F}_V(\cdot \mid x)$ and $\bar{F}_W(\cdot)$ are respectively the conditional CCDF of V and the CCDF of W, respectively. Moreover, there exists a unique stationary distribution $F(y)$ satisfying the following ordinary integro-differential equation:

$$\gamma y \frac{\mathrm{d}F(y)}{\mathrm{d}y} = \int [\lambda(t)\bar{F}_V(y-x \mid x) + \mu \bar{F}_W(y-x)] \frac{\mathrm{d}F(x)}{\mathrm{d}x} \mathrm{d}x. \tag{3}$$

Theorem 2 allows the evolution of the distribution function $F(y,t)$ over time to be traced through a partial integro-differential equation. In addition, the result also provides the stationary distribution, described through an Ordinary Integro-Differential Equation, which describes the system at equilibrium. Again, the proof is available in our technical report cited above.

To corroborate our results, we will present in Section 4 a comparison between the simulation of the Markov process and the theoretical distributions obtained through Theorem 2.

4 Numerical Analysis

In this section, we use our model to explore how key performance indicators depend on various system parameters. We focus on a reference scenario derived from our empirical data, and perform a sensitivity analysis by varying one parameter at a time. This allows us to investigate what-if scenarios and better understand popularity dynamics and influencer competition.

[8] μ_{Leb} denotes the Lebesgue measure

4.1 Metrics and Sensitivity Analysis

We consider a set of five virtual influencers competing for user attention, $\mathcal{I} = \{1, 2, 3, 4, 5\}$. We set $\beta_i = 0.9^{i-1}$ to introduce a systematic variation in their ability to garner user engagement. Note that, although the first influencer is expected to be the most popular one on average, the model allows for the emergence of any influencer as the dominant actor, at least temporarily.

To quantitatively assess the competitive dynamics and performance hierarchy among these influencers, we introduce the following two key metrics:

- **First place probability**, defined as $\pi_1^{[i]} = \mathbb{P}(X^{[i]} \geq X^{[j]}, \forall j \neq i)$, which is the probability that influencer i achieves the highest popularity under stationary conditions. Conceptually, it can also be interpreted as the long-term fraction of time during which the given influencer occupies the top position in the popularity ranking.
- **First place average stay**, denoted with $S_1^{[i]}$, which is the average sojourn time of influencer i in the first place (in stationary conditions), providing insight into the temporal stability of its dominant status.

Inspired by our empirical analysis of data traces, we consider a reference scenario whose parameters are summarized in Table 2. We remark that in our reference scenario we have made a few simplifying assumptions to obtain a simple baseline case: (i) the posting process is the same for all influencers (parameters $\lambda_0, \lambda_1, \phi$ are now independent of i); (ii) the posting process is a simple Poisson process ($\lambda_1 = 0$); (iii) while the mean number of likes obtained by a post is different for each influencer, the coefficient of variation CV of the lognormal distribution is here assumed to be the same for all influencers.

Figure 5 shows, on a log-log scale, the stationary Probability Distribution Function (PDF) of $X^{[i]}$ for the five considered influencers, obtained by simulation and by numerically solving (3). The perfect match between simulation and analysis cross-validates both approaches to obtain the stationary distribution of influencers' popularity.

Starting from the baseline case, Figure 6 explores the impact of popularity decay rate γ on $\pi_1^{[i]}$ (left plot) and $S_1^{[i]}$ (right plot). As expected, as γ diminishes ($1/\gamma$ increases), the top influencer tends to monopolize users' attention, since its popularity decays slower, reflecting the cumulative effect of a large number of posts, which tends to concentrate around the mean (determined by intrinsic ability β_i). The opposite is true for large γ (small $1/\gamma$), where we observe the opposite regime in which all influencers demonstrate comparable probabilities of achieving dominance, as popularity dynamics in this case are principally governed by the stochastic success of individual posts.

Analogous considerations can be done when we fix γ and change the coefficient of variation of the like distribution (see Figure 7). Naturally, concurrent variation of both γ and CV relative to our baseline would produce a compound effect, to be again interpreted through the lens of stochastic versus deterministic behavior (i.e., concentration around the mean) in the popularity dynamics.

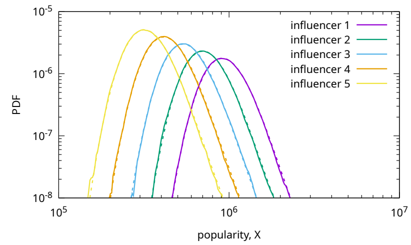


Fig. 5: Popularity distribution of five influencers in the baseline scenario. Comparison between simulation results (solid) and analytical results (dashed lines).

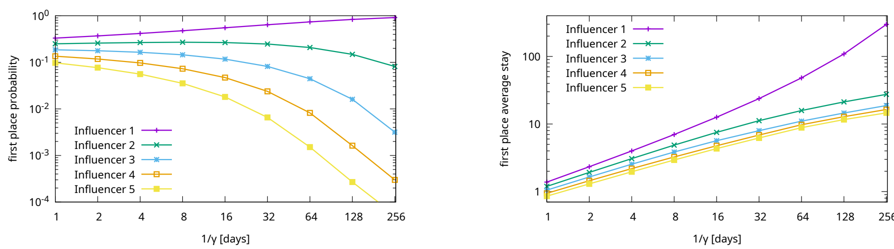


Fig. 6: Sensitivity analysis with respect to popularity decay rate γ. Impact on first place probability (left plot) and first place average stay (right plot).

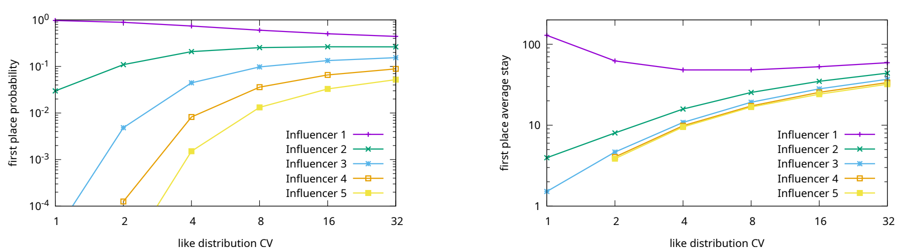


Fig. 7: Sensitivity analysis with respect to the coefficient of variation CV of the like distribution. Impact on first place probability (left plot) and first place average stay (right plot).

We emphasize that a proper notion of fairness among influencers is required to assess whether a given system performs better than another, but a formal definition of this notion extends beyond the scope of the present investigation.

Next, in Figure 8, we consider the impact of exponent θ, describing the (sub-linear) growth of the number of likes collected by a post as a function of the influencer's current popularity. As expected, larger values of θ (but recall that we need $\theta < 1$ for the system to be stable) amplify the disparity among influencers, underscoring the critical role of the platform's engagement mechanism(s) in determining their relative success.

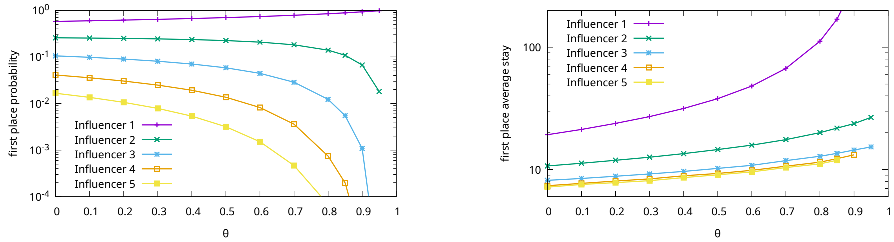


Fig. 8: Sensitivity analysis with respect to exponent θ. Impact on first place probability (left plot) and first place average stay (right plot).

4.2 Impact of Non-Poisson Post Arrival Process

Here, we examine how the burstiness of the post arrival process affects our metrics with respect to the baseline scenario, wherein posts are generated according to a simple Poisson arrival process of fixed intensity 4 posts/day. Recall that in our model we consider an inhomogeneous Poisson process for each influencer i, with intensity $\lambda(t \mid X^{[i]}(t^-) = x) = \lambda_{0,i} + \lambda_{1,i} x^{\phi_i}$, which increases the probability of new post emissions during periods of higher influencer popularity, a pattern consistently observed in empirical data.

To simplify the exploration of the parameter space, we proceed as follows: (i) we establish the constant term of the posting rate $\lambda_0 = 1$, equal for all influencers; (ii) for ϕ, we consider $\{0, 0.1, 0.2, 0.3\}$ as possible values, equal for all influencers; (iii) we systematically vary λ_1, setting it equal for all influencers.

Figure 9 shows the results of this experiment for $\phi = 0.2$, reporting the first place probability (solid lines, left y axes) and the obtained average posting rate (dashed lines, right y axes). As expected, larger values of λ_1 produce larger values of average posting rate, amplifying discrepancies among the five influencers. To isolate the effect of non-homogeneous post arrival rate, while ensuring a fair competition among the influencers, we implemented the following methodology: for each considered value of ϕ, and for each influencer, we numerically determined the value of λ_1 that yields an average posting rate of 4 posts/day (as in the baseline scenario).

In Figure 9, such values of λ_1, in the case of $\phi = 0.2$, are those at which dashed lines intersect the horizontal dotted line plotted at $y = 4$. Subsequently, we recalculated the first place probability in a unique scenario utilizing these calibrated λ_1, with results in Table 3. Note that the first column ($\phi = 0$) corresponds to the case of Poisson post arrival rate. We observe that, as we increase ϕ, thereby enhancing the burstiness of the post arrival process, discrepancies among influencers tend to reduce.

5 Discussion and Concluding Remarks

Our work combines theoretical modeling of popularity dynamics with empirical data from Facebook, providing insights into the complex interaction of fac-

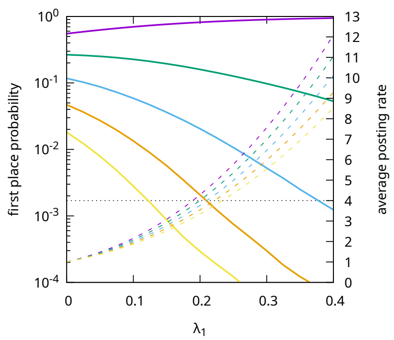


Fig. 9: Sensitivity analysis with respect to λ_1 , for fixed $\lambda_0 = 1$, $\phi = 0.2$.

Table 3: First place probability with adapted λ_1, so that the average posting rate is constant (4 posts/day) for all influencers.

Influencer	$\phi = 0$	$\phi = 0.1$	$\phi = 0.2$	$\phi = 0.3$
1	0.738	0.701	0.651	0.597
2	0.208	0.226	0.253	0.264
3	0.045	0.058	0.073	0.099
4	0.008	0.012	0.019	0.032
5	0.002	0.002	0.004	0.009

tors like individual activity, popularity decay, content attractiveness, and the platform's role in content visibility. Besides deriving ergodicity conditions, we conducted a sensitivity analysis to explore the impact of system parameters on possible outcomes.

In principle, an ideal system should provide each influencer with a fair opportunity to gain attention, commensurate with their intrinsic merit. In this context, one should aim to avoid two problematic extremes: (i) a situation in which the most attractive influencer monopolizes attention, leaving too little visibility for others; and (ii) a condition in which influencers are roughly equally likely to gain primacy, irrespective of their intrinsic merit or capabilities. Fairness, therefore, involves not only fostering an environment of discussion and information exchange in which each influencer has a chance to gain public attention commensurate with merit, but also ensuring that the attention amplification induced by recommendation systems and platform algorithms does not distort competition. Finding an optimal balance between the above two extremes involves complex ethical, technological, sociological, and political issues that are beyond the scope of this analysis. However, the proposed model, by offering a quantitative tool for examining how various factors influence the considered metrics, allows one to explore "what-if" scenarios, yielding valuable insights into popularity dynamics and competitive processes within online social networks.

Subsequent research will extend the analysis to other platforms, develop a framework for influencer competition, and explore control strategies to foster equity in competitive environments.

Acknowledgments. This research has been partially supported by the European Union – Next Generation EU by the Spoke 1 "FutureHPC & BigData" of ICSC - Centro Nazionale di Ricerca in High-Performance-Computing, Big Data and Quantum Computing and by Mission 4, Component 1, under the PRIN project TECHIE: "A control and network-based approach for fostering the adoption of new technologies in the ecological transition" Cod. 2022KPHA24 CUP Master: D53D23001320006, CUP: B53D23002760006.

Disclosure of Interests. Authors do not have any competing interests to declare.

References

1. Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J.: Everyone's an influencer: quantifying influence on twitter. In: ACM WSDM '11. New York, NY, USA (2011)
2. Bass, F.M.: A new product growth for model consumer durables. Management Science **15**(5), 215–227 (1969)
3. Bessi, A.: On the statistical properties of viral misinformation in online social media. Physica A: Statistical Mechanics and its Applications **469**, 459–470 (2017)
4. Castaldo, M.: Attention dynamics on YouTube : conceptual models, temporal analysisof engagement metrics, fake views. Theses, Université Grenoble Alpes (2022)
5. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.: Measuring user influence in twitter: The million follower fallacy. Proceedings of the International AAAI Conference on Web and Social Media **4**(1), 10–17 (May 2010)
6. Collet, F., Pra, P.D., Sartori, E.: A simple mean field model for social interactions: Dynamics, fluctuations, criticality. Journal of Statistical Physics **139**(5) (2010)
7. Covington, P., Adams, J., Sargin, E.: Deep neural networks for youtube recommendations. In: Proc. of the 10th ACM Conference on Recommender Systems (2016)
8. Crane, R., Sornette, D.: Robust dynamic classes revealed by measuring the response function of a social system. Proceedings of the National Academy of Sciences **105**(41), 15649–15653 (2008)
9. De Choudhury, M.: Modeling and predicting group activity over time in online social media. In: Proceedings of the 20th ACM Conference on Hypertext and Hypermedia. p. 349–350. HT '09, New York, NY, USA (2009)
10. Doshi, R., Ramesh, A., Rao, S.: Modeling influencer marketing campaigns in social networks. IEEE Trans. on Computational Social Systems **10**(1), 322–334 (2023)
11. Friedkin, N., Johnsen, E.: Social influence networks and opinion change. Advances in Group Processes **16** (01 1999)
12. Hilgartner, S., Bosk, C.L.: The rise and fall of social problems: A public arenas model. American Journal of Sociology **94**(7), 53–78 (1988)
13. Jiang, R., Chiappa, S., Lattimore, T., György, A., Kohli, P.: Degenerate feedback loops in recommender systems. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. p. 383–390. AIES '19 (2019)
14. Lorenz-Spreen, P., Mønsted, B.M., Hövel, P., Lehmann, S.: Accelerating dynamics of collective attention. Nature Communications **10**(1), 1759 (2019)
15. Meyn, S.P., Tweedie, R.L.: Markov chains and stochastic stability. Communications and Control Engineering, Springer, London, England, 1993 edn. (Dec 2012)
16. Muchnik, L., Aral, S., Taylor, S.: Social influence bias: A randomized experiment. Science **341**(6146), 647–651 (2013)
17. Ravazzi, C., Dabbene, F., Lagoa, C., Proskurnikov, A.V.: Learning hidden influences in large-scale dynamical social networks: A data-driven sparsity-based approach. IEEE Control Systems Magazine **41**(5), 61–103 (2021)
18. Richier, C., Altman, E., Elazouzi, R., Jimenez, T., Linares, G., Portilla, Y.: Bio-inspired models for characterizing youtube viewcout. In: ASONAM (2014)
19. Vassio, L., Garetto, M., Leonardi, E., Chiasserini, C.F.: Mining and modelling temporal dynamics of followers' engagement on online social networks. Social Network Analysis and Mining **12**(1), 96 (2022)
20. Weng, L., Flammini, A., Vespignani, A., Menczer, F.: Competition among memes in a world with limited attention. Scientific Reports **2**(1), 335 (2012)