

Collective Attention to Social Media Evolves According to Diffusion Models

Christian Bauckhage^{1,3}
¹University of Bonn
53113 Bonn, Germany

Kristian Kersting^{2,3}
²TU Dortmund University
44221 Dortmund, Germany

Bashir Rastegarpanah
³Fraunhofer IAIS
53754, St. Augustin, Germany

ABSTRACT

We investigate patterns of adoption of 175 social media services and Web businesses using data from Google Trends. For each service, we collect aggregated search frequencies from 45 countries as well as global averages. This results in more than 8.000 time series which we analyze using economic diffusion models. The models are found to provide accurate and statistically significant fits to the data and show that collective attention to social media grows and subsides in a highly regular manner. Regularities persist across regions, cultures, and topics and thus hint at general mechanisms that govern the adoption of Web-based services.

Categories and Subject Descriptors

H.3.5 [Online Information Services]: Web-based services

Keywords

social media services, collective attention, trend prediction

1. INTRODUCTION

Understanding the dynamics of collective human attention has been called a key scientific challenge for the information age [4]. We address a specific aspect of this problem and mine query logs for common trends and shared characteristics. Our focus is on Google Trends data which summarize the evolution of global and regional interests in *social media services* and *Web businesses* and we explore to what extent the general dynamics of collective attention apparent from these data can be modeled mathematically.

Given a comprehensive empirical basis, we perform trend analysis using economic diffusion models and find them to be in excellent agreement with the data. In particular, we find that collective attention to social media as evident from search frequencies evolves according to notably regular patterns. Although microscopic behaviors may be chaotic, general trends apparent in these data typically show simple and

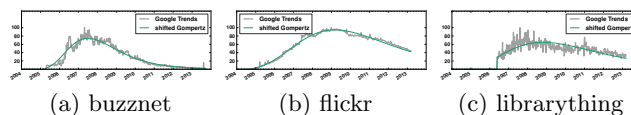


Figure 1: Time series summarizing how worldwide searches for social media services evolve over time; a fitted diffusion model accounts well for the general trends of growing and declining interest.

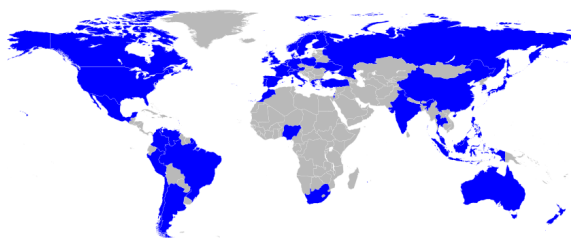


Figure 2: 45 countries considered in this study

highly regular dynamics of growth and decline which persist across regions, cultures, and linguistic backgrounds.

2. METHODOLOGY

Google Trends provides statistics about queries users submitted to Google's search engine. Beginning from January 1st 2004, weekly aggregated statistics are available in form of global averages but can be narrowed down to statistics for individual countries. Here, we analyze global and regional data related to queries such as *ebay*, *facebook*, or *youtube* that indicate interest in social media services or Web-based businesses. In total, we consider data as to 175 services that were chosen according to penetration and profile. They include social networks, photo- and video sharing sites, music streaming services, virtual hangouts, (micro-)blogging services, and online retailers, trading platforms, as well as social games providers. The 45 countries considered were selected according to population size, Internet penetration, and availability of query logs. Our sample covers various regions, cultures, and official languages. As we also retrieve corresponding global search activities, our empirical basis consists of more than 8.000 data sets. For data harvesting, processing, and analysis, we follow the protocol in [2].

Visual inspection of our data reveals noticeably common patterns: macroscopically, collective interest in a service

Table 1: Goodness of fit w.r.t. regions of the world

region	shifted Gompertz		Bass		Weibull	
	$\langle p \rangle$	$p > 5\%$	$\langle p \rangle$	$p > 5\%$	$\langle p \rangle$	$p > 5\%$
Africa	0.61	68%	0.55	62%	0.50	57%
Asia	0.57	63%	0.49	54%	0.48	53%
Australia	0.66	70%	0.53	59%	0.50	58%
Europe	0.59	65%	0.48	51%	0.56	54%
N-America	0.54	57%	0.44	50%	0.39	44%
S-America	0.65	71%	0.54	59%	0.55	62%
worldwide	0.59	64%	0.50	55%	0.47	53%

Table 2: Goodness of fit w.r.t. languages of the world

language	shifted Gompert		Bass		Weibull	
	$\langle p \rangle$	$p > 5\%$	$\langle p \rangle$	$p > 5\%$	$\langle p \rangle$	$p > 5\%$
English	0.55	58%	0.44	49%	0.39	45%
Spanish	0.63	68%	0.52	56%	0.54	60%
Portuguese	0.60	67%	0.50	56%	0.47	51%
Russian	0.68	76%	0.58	66%	0.69	76%
French	0.55	60%	0.46	51%	0.39	45%
German	0.58	64%	0.47	52%	0.47	54%
Chinese	0.50	52%	0.42	46%	0.43	47%
Japanese	0.42	52%	0.38	44%	0.31	38%
Hindi	0.57	64%	0.47	54%	0.48	52%
average	0.57	62%	0.47	52%	0.45	51%

shows a phase of accelerated growth followed by saturation and prolonged decline. Skewed time series like these often occur in economics where they indicate buying behaviors or adoption rates and are studied using *diffusion models*. We therefore investigate if diffusion models fit our data. In particular, we examine the Bass distribution [1], the shifted Gompertz distribution [3], and the Weibull distribution [2]. For a discussion as to the socio-mathematical underpinnings of these models, we refer to [2].

3. RESULTS

Table 1 presents Goodness-of-Fit (GoF) results for our models in terms of p -values obtained from χ^2 -tests. At a significance level of 5%, we find the shifted Gompertz to provide accurate fits for the majority of our data and its average p -values (higher is better) significantly exceed 0.5. This holds for fits to data which reflect worldwide interests as well as for continent specific data. Table 2 lists GoF results w.r.t. major languages. These results mimic those in Tab. 1; quality and significance of fits are comparable and the shifted Gompertz again provides the most accurate explanation. These results are interesting for they suggest that *the dynamics of collective attention apparent from search frequency data can be accurately described in terms of diffusion models*. Moreover, they indicate that *collective attention to social media evolves similarly and independent of regions of origin or cultural backgrounds of crowds of Web users*.

An advantage of two-parameter diffusion models is that they facilitate *visual analytics*. Once a model is fit to a time series, its parameters $[\theta_1, \theta_2]$ provide a non-linear 2D embedding of the data. Figure 3 displays such embeddings for all our data. In each case, *amazon*, a business that continues to attract user interest, marks an extreme location. Similarly extreme extreme locations are occupied by *craigslist* and *ebay*, two platforms that were launched in the 1990s and reached global

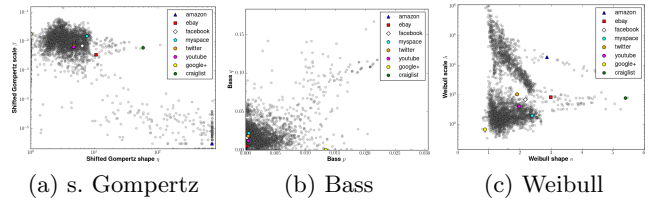


Figure 3: Non-linear, 2D embeddings of over 8,000 search frequency time series into the parameters spaces of the diffusion models considered here.

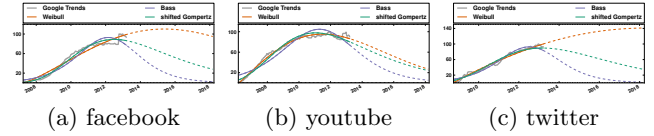


Figure 4: Predictions of future collective interest in exemplary social media services. Note that these plots predict relative search frequencies w.r.t. the maximum interest so far which is scaled to 100.

peak popularity around 2008. The coordinates of *google+*, a service whose search frequencies spiked after its launch in 2011, reside at opposite extremes. For each model, however, most time series are found in a giant cluster. The existence of these giant clusters of almost 90% of all time series tested is arguably our most important finding. Irrespective of the diffusion model used for analysis and regardless of region or language, most of our time series show similar behavior and the shape and scale parameters of fitted diffusion models indicate that *individual social media services seem to be able to attract increasing collective attention for a period of 4 to 6 years before user interest inevitably begins to subside*.

Finally, we apply our models to predict future evolutions of global collective interest. Figure 4 shows predictions for three of today’s prominent social media services. Here, the shifted Gompertz marks a middle ground. For *facebook* it predicts collective interest to reduce to 50% of its current intensity by 2017. While this may seem improbable from today’s point of view, we note again that the vast majority of the 175 services analyzed here show characteristic cycles of growth and decline. Given the data that is publicly available as of this writing, collective attention to *facebook* so far seems to follow the same pattern.

4. REFERENCES

- [1] F. Bass. A New Product Growth Model for Consumer Durables. *Management Science*, 15(5):215–227, 1969.
- [2] C. Bauckhage, K. Kersting, and F. Hadiji. Mathematical Models of Fads Explain the Temporal Dynamics of Internet Memes. In *Proc. ICWSM. AAAI*, 2013.
- [3] A. Bemmaor. Modeling the Diffusion of New Durable Goods : Word-of-mouth Effect Versus Consumer Heterogeneity. In G. Laurent, G. Lilien, and B. Pras, editors, *Research Traditions in Marketing*, pages 201–229. Springer, 1994.
- [4] F. Wu and B. Huberman. Novelty and Collective Attention. *PNAS*, 104(45):17599–17601, 2007.