Network and Revenue of the *Clube Hurb* Affiliate Marketing Program: A Story of Two Tales

Lucas L. Rolim
Hurb.com
Rio de Janeiro, Brazil
lucas.rolim@hurb.com

Jefferson E. Simões
Applied Informatics Department
UNIRIO
Rio de Janeiro, Brazil
jefferson.simoes@uniriotec.br

Daniel R. Figueiredo
Systems Engineering and Computer Science
Federal University of Rio de Janeiro (UFRJ)
Rio de Janeiro, Brazil
daniel@cos.ufrj.br

Abstract—Individuals in affiliate marketing programs sign up with companies to promote or sell their products in independent venues and channels, receiving compensations for their actions. While a component of the e-commerce ecosystem for over a decade, affiliate marketing is increasingly being adopted by companies given its promises of boosting revenue at low investment costs. This work analyzes *Clube Hurb*, the affiliate marketing program of Hurb.com the largest online travel agency in Brazil. The analysis reveals the fragility of social network growth (very low virality coefficient) along with the strength of social referrals. It also reveals that almost all revenue generated by the program comes from a small set of affiliates, a property that has sustained over time. Indeed, great disparities are characterized by heavy-tailed distributions in statistics concerning both the network and revenue structure. Thus, while most affiliates play no effective role, a few are instrumental in keeping the program profitable. Such findings indicate that traditional average-based performance metrics can be flawed when assessing the success of such programs.

Index Terms—Travel industry, affiliate marketing, affiliate network, network analysis, revenue analysis

I. INTRODUCTION

The widespread adoption of e-commerce has transformed various markets and continues to grow both in size and scope, in particular with the emergence of online marketplaces and alternative payment systems [1]. The role of marketing and advertisement has also been transformed, in part due to services provided by online social networks (such as Facebook) and search engines (such as Google).

Within the context of online social networks, *viral marketing* has emerged as promising and effective technique for product and service advertisement [2]–[4]. A relatively less studied approach is *affiliate marketing*, which is being increasingly adopted by companies due to its promises of boosting revenue at a relatively low investment cost [5]–[7].

In affiliate marketing, a person (the affiliate) signs an agreement with a company to promote and/or sell products and services in venues independent from those of the company, such as websites, blogs and online social networks. In return, the company pays a commission to the affiliate for different kinds of actions such as selling a product or attracting customers to the company's website.

In order to increase the number of affiliates, a company often compensates an affiliate that brings other affiliates to their program through referrals. This naturally leads to an *affiliate network*, a tree where nodes are affiliates and the parent of a node is the affiliate that referred that node to the program. Usually affiliates can often enroll directly into the program, thus becoming a root of its own tree.

In viral marketing, the online social network structure plays a key role on the success of a marketing campaign, in particular due to the presence of individuals with large degrees (i.e., hubs). In such contexts, the *reproduction number* (i.e., average out-degree) is often taken as a metric to quantify the success of a campaign [4], [8]–[10]. Do affiliate networks also have such properties? Are hubs important in affiliate networks?

This work contributes with an extensive analysis of *Clube Hurb*. Launched in April 2018, the affiliate marketing program of Hurb.com accumulated over 186,000 affiliates by January 2020, who generated significant revenue through sales, and continues to grow to this day. This program's data was collected and analyzed, leading to the following main findings and contributions: extensive structural analysis about affiliates referral dynamics, network's performance analysis in terms of revenue generation, and relationship between structural and performance characteristics on a real large-scale affiliate network.

The above empirical findings shed light on the heterogeneity of an affiliate marketing program, where many network and revenue metrics exhibit high variance and heavy tail distributions. This gives rise to a story of two tales, where the first and superficial tale, told by classic online social network metrics, fails to assess the success of *Clube Hurb*, whereas the second tale, which complements them more deep-looking metrics, is able to better understand its success. For instance, while referrals in *Clube Hurb* have failed to attract a large number of affiliates, it has attracted affiliates that generate significantly more revenue (on average). Our findings can be leveraged to create incentive mechanisms within *Clube Hurb*, to accelerate its growth and boost its revenue.

The remainder of this paper is organized as follows. Section II discusses related work on affiliate marketing. Details

of *Clube Hurb* and the dataset are provided in Section III. Section IV shows the referral and revenue analysis. A discussion of the empirical findings and its business implications in provided in Section V, and a brief summary in Section VI.

II. RELATED WORK

The widespread expansion of e-commerce worldwide has profoundly transformed several market segments and impacted from retailers to consumers. Along with it, marketing and advertisement have also been profoundly transformed over the past two decades. New marketing paradigms such as viral marketing and affiliate marketing have emerged and proven to be effective, playing an important role in the online ecosystem [2]–[5], [7], [11].

Intuitively, *viral marketing* is the online version of word-of-mouth: each person spreads information concerning a product or a service that might be relevant to others according to their own judgment or opinion. Moreover, the online community has grown exponentially in size, along with the number of channels for interaction such as blogs and social media. Under this scenario, viral marketing has been hailed to have an enormous potential, yet to be fully understood and realized [2], [3], [8], [12]. Problems such as how to start a successful viral marketing campaign has received much attention from academia and industry [9], [10].

Related but distinct is the concept of *affiliate marketing*. In affiliate marketing, individuals sign an agreement with a company to promote and sell its products and services in independent platforms, such as websites, blogs and channels owned or controlled by the affiliate [5], [7], [11]. Under such agreements, the affiliate receives direct compensation from the company for performing actions such as selling products, attracting customers to the company's website, or referring new affiliates. While affiliate marketing is a fast growing trend among companies, it has been relatively less studied both in academia and industry [6], [13].

Two fundamental aspects of affiliate marketing are determining the compensation value for the various actions of the affiliates, as well as the disclosures of a company's internal information to the affiliates (which enables them to promote and/or sell products and services). The careful execution of both tasks is critical to the success of affiliate marketing [13].

Within the travel industry, recent studies have analyzed the qualitative determinants of consumer trust in affiliate marketing networks for tourism companies [11] and built recommendation systems using social network information [14]. However, how to exploit the social network structure and its features to improve sales or boost affiliate growth remains an open question.

III. *Clube Hurb* AND DATASET

Hurb.com is the largest Brazilian online travel agency (OTA) and one of the largest in the world, selling thousands trips per month to hundreds of thousands of travelers. In April 2018, it launched *Clube Hurb*, a public affiliate marketing program to create a novel sales channel, attract new clients, and increase brand visibility. The program offered direct affiliate enrollment via ads on its website and on platforms like Google and Facebook. An affiliate is paid a flat commission rate of 6.5% for any product it sells, though special, higher commission rates apply in specific occasions (e.g. Black Friday). To help its affiliates sell products, *Clube Hurb* provides a platform to create personalized ads that can be shared on social media platforms and posted on blogs and websites.

Clube Hurb also allows affiliates to refer the program to other potential affiliates. When an person enrolls in the program through a referral, the referee earns 15% of the commission received from every sale the new affiliate makes. This commission is paid only for direct referrals, thus excluding referrals of referrals and so on.

Note that the program rewards an affiliate by the performance of their referrals, and not when a referral joins the program. Thus, every referral that becomes an affiliate is a potential continuous source of income for the referee. This provides strong incentives for an affiliate to bring in new affiliates. On the other hand, direct enrollment bears no cost to the new affiliate, as their commission on sales is not shared by a parent affiliate.

The dataset used in this work considers all affiliates registered in *Clube Hurb* between the program's inception, in April 2018, and January 2020. Various kinds of data were collected for this study, including data on affiliates that joined through referrals, their sales and earnings, and associated company revenue.

IV. NETWORK AND REVENUE ANALYSIS

The process of enrollment in the affiliate program through referrals naturally induces a directed social network. In this network, every affiliate is a node, and a directed edge from u to v exists if v enrolled in the affiliate program through a referral from u. The affiliate network of *Clube Hurb*, constructed from the collected data, will be henceforth denoted by $G = (V, E)$ where V is the set of nodes (affiliates) and E the set of directed edges (enrolled referrals). Since affiliates may join the program either directly or through a referral, nodes in G have in-degree either 0 and 1, and G has no cycles. Therefore, G is a set of directed trees, i.e., a directed forest[1]. A detailed analysis of the network structure can be found in [15].

TABLE I: Affiliate network features (as of January 2020)

Attribute	Value
Number of nodes	186,248
Number of edges	11,536
Number of connected components	174,712
Number of isolated nodes	169,081
Largest connected component	309
2nd largest connected component	174
Fraction of nodes with a parent	6.19%
Fraction of nodes with no child	96.68%
Fraction of nodes with one child	2.65%

---

[1]More precisely, each connected component of G is an *out-tree* or *arborescence*: a directed tree with edges pointing away from the root.

A. Referral Analysis

While most affiliates essentially do not participate in referrals, it is still interesting to characterize the 11,536 edges of the affiliate network. The average out-degree is 0.06 and median out-degree is 0; both statistics are heavily influenced by the large amount of isolated nodes. However, the distribution is heavy-tailed, and a non-negligible number of affiliates have out-degree larger than 50 (maximum out-degree is 228).

How long does it take for nodes to successfully refer other nodes to the program? Let Δ be the difference (in days) between the time a parent and its referral joined the program (this is measured for every edge in G). The distribution of Δ is shown in Figure 1-L. Note that, despite a relatively low median of 30 days, this distribution exhibits high variance, with 10% of referrals joining the network 150 days after their parent. Interestingly, some affiliates can attract new affiliates even after 500 days of their joining. Thus, while novelty is important in converting referrals, some affiliates can convert referrals even after long being part of the network.

Can referrals boost the revenue generated by an affiliate? Figure 1-M shows a scatter plot of out-degree and generated revenue for each affiliate. There is no clear visual correlation between the two statistics, and indeed, the measured Pearson correlation coefficient is equal to 0.03. While classic network analytics take a node's degree as a measure of its value or importance in the network, in our network this ultimately does not translate to effective value (revenue).

Regarding the network formation process, we can consider the growth of the affiliate network to be modeled as a network epidemic process, with "infection" being analogous to membership to the program'. In this model, the average out-degree indicates the expected number of new "infections" per infected individual, a quantity known as the *reproduction rate* of the epidemic process [4]. Figure 1-R shows the accumulated reproduction number of *Clube Hurb* over time, indicating an initial growth when referrals were introduced before converging to around 0.06, a very low number. While reproduction rate plays a key role on viral marketing and other diffusion processes on social networks, its role on this affiliate network is marginal and steady.

B. Revenue Analysis

It is also interesting to understand how different affiliates contribute to the revenue. A common metric is the average performance (revenue) of certain groups of users (affiliates). This is known as *user lifetime value* (LTV).

Table II presents the LTV for various groups of affiliates over the lifetime of the program. Note that the baseline LTV of an affiliate with no known features is 70.59. Yet, conditioning on method of affiliation, we can see that the LTV of affiliates with a parent (i.e., who joined through a referral) is 175.79 and the LTV of affiliates without a parent (i.e. who joined on their own) is 60.32. Thus, having a parent boosts the LTV of an affiliate by a factor of almost 2.5 relative to the baseline, an indication that the referral program contributes to attracting individuals which generate more average revenue.

TABLE II: Average revenue per affiliate (LTV)

Affiliate group	LTV (R$)
All affiliates	70.59
Affiliates with parents	175.79
Affiliates without parents	60.32
Affiliates with parents, 1+ sales	1925.84
Affiliates without parents, 1+ sales	2018.03

It can be seemingly inferred that having a parent directly influences a node's generated revenue, the actual impact in terms of revenue comes through having or not a sale. Table II also presents the LTV for affiliates that make at least one sale, split between affiliates with or without a parent. Note that, relative to the baseline, making at least one sale increases the LTV of the group by a factor of 11. Moreover, conditionally on this, there is little difference between having a parent or not. Thus, the economic success of the affiliate program strongly depends on the affiliate making their first sale, which in turn is positively correlated to having a parent in the network.

Figure 2-L shows the LTV of the affiliate population over time. There is no clear trend in the evolution of LTV, except for a steady decline over the last 6 months of data. Interestingly, the accumulated fraction of affiliates that have not generated any revenue, depicted in Figure 2-R, has remained nearly constant, oscillating around 0.99, throughout the whole data collection period. Thus, while the affiliate program grows steadily over time, the fraction of affiliates with no revenue remains high and constant. This indicates that, in order to understand the trends in revenue, it is important to take a closer look at the distribution of revenue generated by users.

To do this, we will now focus on the inequality of generated revenue, as indicated by the Lorenz curve [16]. The Lorenz curve was originally proposed and is traditionally used to measure income inequality, and denotes the fraction of income accumulated by the fraction of population with lowest income. The Lorenz curve is also related to the Gini Coefficient, one of the most well-known measures of income inequality. More precisely, the Gini Coefficient is the relative area between the Lorenz curve and the "curve of perfect equality", with higher values indicating higher levels of inequality.

In our context, as we wish to measure inequality in generated revenue, our population will consist of all affiliates in the program that have made at least one sale[2] and the variable of interest is individual generated value. The resulting curve can be seen in Figure 3-L. Note that, even within this group, inequality is still very high, with a measured Gini coefficient of 0.79. This indicates that the economic success of *Clube Hurb* highly depends on the performance of a very small set of affiliates. For instance, around 75% of all revenue is generated by the top 10% of affiliates with at least one sale, which, in turn, represent only 0.052% of the entire program population.

---

[2]Since 99.48% of affiliates generate no revenue, they would prevent the Lorenz curve from revealing patterns among the remaining affiliates.
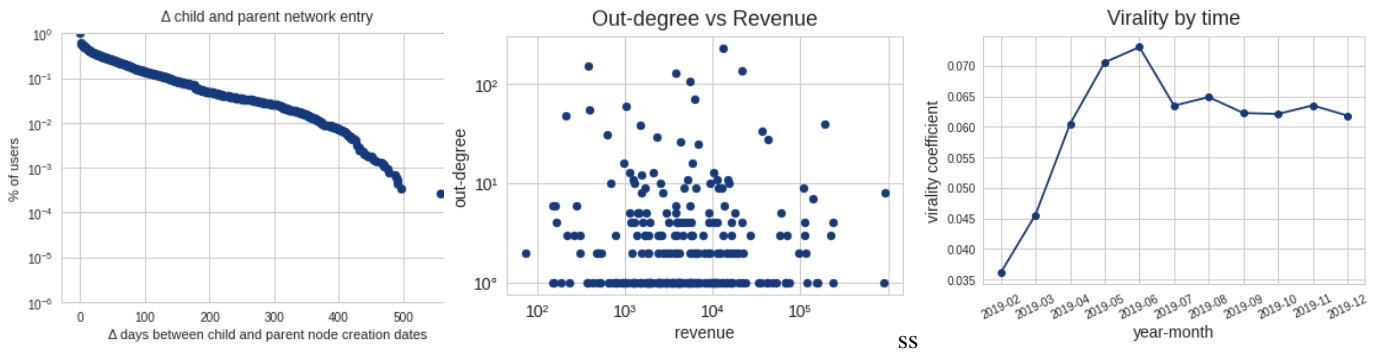
Fig. 1: Complementary cumulative distribution of Δ (difference between join time of parent and referral), in days (left); scatter plot between revenue and out-degree of affiliates (middle); Accumulated reproduction rate or average out-degree (right).
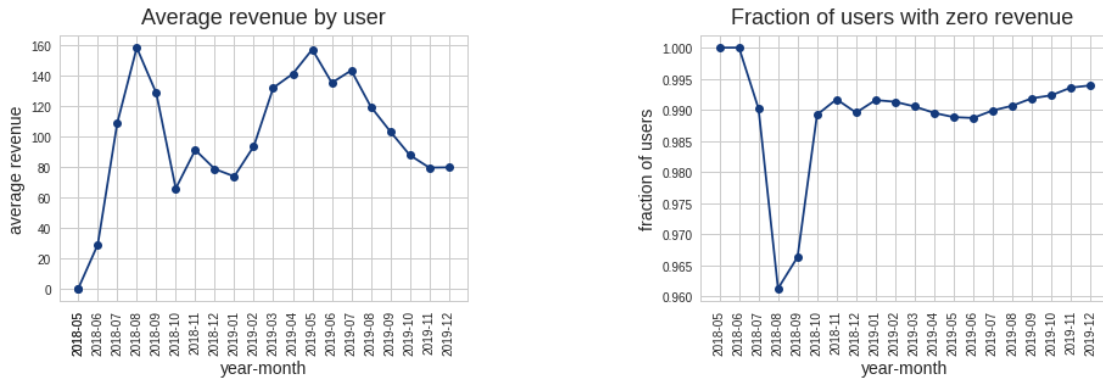

Fig. 2: LTV (average revenue) of affiliates over time (left); Fraction of affiliates with zero revenue over time (right).
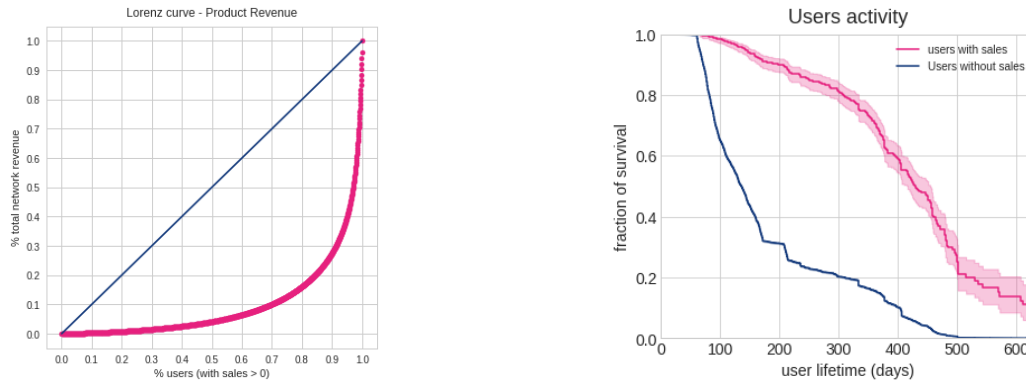

Fig. 3: Lorenz curve for revenue generated by affiliates with at least one sale, with "perfect equality" shown in blue (left); Kaplan–Meier survival curves for all affiliates and for affiliates with at least one sale (right).

C. Activity lifetime

An important expected behavior of affiliates is their engagement with the program. Affiliates are able to engage by generating personalized ads and publicizing them in online platforms, in order to convert sales and earn commissions. While this is difficult to track, there is a reasonable proxy for user engagement via logins to the affiliate program. How often do affiliates login and what is their expected level of activity?

A common methodology to analyze life expectancy is the Kaplan–Meier curve, which indicates the fraction of users (within a population) with lifetimes larger than a given period. In the context of affiliate programs, we consider that a "birth" occurs when the user registers in the program, and a "death" occurs when the user has remained for some period of time without logging in. In the forthcoming analysis, we have fixed the death threshold at 60 consecutive days without a login.

Figure 3-R shows the Kaplan–Meier curves for all affiliates and only for affiliates with at least one sale. Note that only 20% of all affiliates survive more than 300 days. However, among affiliates with at least one sale, 80% survive more than

300 days. Having a sale also provides a roughly threefold increase in median lifetime, from 130 to 420 days. Thus, having a single sale not only significantly improves user LTV (as shown in Table II) but also improves user lifetime.

V. DISCUSSION AND BUSINESS IMPLICATIONS

Online social network services are often evaluated according to aggregate metrics, such as the reproduction number, user lifetime, and fraction of new users. From this perspective, the affiliate network of *Clube Hurb* is a failure. Through its first 12 months, its referral program has failed to attract a significant number of new users, despite the commission paid to referees. Median lifetime of users is only 150 days. While hubs are important in classic online social network services, affiliates in *Clube Hurb* show no correlation between degree and revenue. However, affiliates that join through referrals do perform better than others, a clear indication of the potential of referrals.

Of course, online services are also evaluated according to the revenue generated over time, in particular revenue growth and average user revenue. In a sense, this more fundamental assessment is crucial for a service to become profitable and thus continue to be offered. From this perspective, *Clube Hurb* is a success. The revenue generated by the affiliate program has grown steadily for the last 12 months and shows no sign of slowing down. Clearly, some affiliates reach a client base that is not otherwise accessible and lead them to purchases. However, revenue is far from homogeneous among affiliates, where the vast majority do not generate any revenue and a very small fraction is responsible for almost all revenue generated. Interestingly, this fraction is consistent over time.

Thus, *Clube Hurb* is a story of two tales. It thrives as a successful online service from the perspective of revenue it generates, but it fails from the perspective of an online social network and from the perspective of the majority of affiliates, that do not benefit from commissions (since they do not generate revenue). Can it do better?

A. Interventions

The analysis of *Clube Hurb* reveals that the program has a diverse set of affiliates, leading to heavy tailed distributions of different network and revenue statistics. However, the analysis also revealed the strength of the conversion funnel with respect to having a parent and, specially, making the first sale. In both cases, the LTV is much larger and lifetime is much longer (see Table II and Figure 3.

Thus, *Clube Hurb* should encourage its affiliates to acquire both characteristics. For example, affiliates should be encouraged to make referrals, possibly by a revision on the compensation scheme, granting an immediate commission to the referrer as soon as a new referral joins the network. Moreover, affiliates should be encouraged to make their first sale, for instance by helping each affiliate create personalized ads, tailored to their individual social circle.

Such interventions could boost user engagement and revenue generation, leading to a more sustainable and profitable program, and the design and analysis of such intervention strategy is the most important line of future work and research.

VI. CONCLUSION

This work provided a detailed analysis of the *Clube Hurb* affiliate program, from both network and revenue perspectives, assessing the metrics over time and their mutual influence. While the program is growing in number of affiliates and generated revenue, its affiliates are heterogeneous, with various metrics exhibiting high variance and heavy-tailed distribution. Thus, classic performance metrics used to assess OSN services are not suited for *Clube Hurb*.

However, the analysis also indicated that affiliates that enroll through referrals and affiliates that make at least one sale generate more revenue. Thus, engaging these characteristics is important to accelerate the growth and boost the revenue of the program. Designing incentive mechanisms to drive these characteristics in new and current affiliates is theme for future research.

REFERENCES

[1] K. C. Laudon and C. G. Traver, *E-Commerce 2019: Business, Technology and Society*, 15th ed. Pearson, 2019.

[2] J. Y. Ho and M. Dempsey, "Viral marketing: Motivations to forward online content," *Journal of Business research*, vol. 63, no. 9-10, pp. 1000–1006, 2010.

[3] T. Reichstein and I. Brusch, "The decision-making process in viral marketing—a review and suggestions for further research," *Psychology & Marketing*, vol. 36, no. 11, pp. 1062–1081, 2019.

[4] D. Watts and J. Peretti, "Viral marketing for the real world," *Harvard Business Review (HBR)*, May 2007.

[5] D. L. Duffy, "Affiliate marketing and its impact on e-commerce," *Journal of Consumer Marketing*, vol. 22, no. 3, pp. 161–163, 2005.

[6] Y. K. Dwivedi, N. P. Rana, and M. A. A. Alryalat, "Affiliate marketing: An overview and analysis of emerging literature," *The Marketing Review*, vol. 17, no. 1, pp. 33–50, 2017.

[7] M. Ivkovic and D. Milanov, "Affiliate internet marketing: Concept and application analysis," in *IEEE International Conference on Education and Management Technology*, 2010, pp. 319–323.

[8] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB)*, vol. 1, no. 1, 2007.

[9] O. Hinz, B. Skiera, C. Barrot, and J. U. Becker, "Seeding strategies for viral marketing: An empirical comparison," *Journal of Marketing*, vol. 75, no. 6, pp. 55–71, 2011.

[10] H. T. Nguyen, M. T. Thai, and T. N. Dinh, "Stop-and-stare: Optimal sampling algorithms for viral marketing in billion-scale networks," in *ACM International Conference on Management of Data (SIGMOD)*, 2016, pp. 695–710.

[11] N. Gregori, R. Daniele, and L. Altinay, "Affiliate marketing in tourism: determinants of consumer trust," *Journal of Travel Research*, vol. 53, no. 2, pp. 196–210, 2014.

[12] S. Goel, D. J. Watts, and D. G. Goldstein, "The structure of online diffusion networks," in *ACM Conference on Electronic Commerce*, 2012, pp. 623–638.

[13] S. A. Suryanarayana, D. Sarne, and S. Kraus, "Information disclosure and partner management in affiliate marketing," in *First International Conference on Distributed Artificial Intelligence*, 2019, pp. 1–8.

[14] A. Menk, L. Sebastia, and R. Ferreira, "Recommendation systems for tourism based on social networks: A survey," *arXiv preprint arXiv:1903.12099*, 2019.

[15] L. L. Rolim, J. E. Simões, and D. R. Figueiredo, "Network and revenue analysis of an affiliate marketing program in the travel industry," in *Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*, 2020, pp. 120–131.

[16] M. O. Lorenz, "Methods of measuring the concentration of wealth," *Publications of the American statistical association*, vol. 9, no. 70, pp. 209–219, 1905.