Activation Games in Online Dating Platforms

Eitan Altman, Francesco de Pellegrini, Huijuan Wang

Activation Games in Online Dating Platforms

Eitan Altman⋆, Francesco De Pellegrini⋄ and Huijuan Wang†

***Abstract*—In this paper we describe a model for the activation level of users in online dating platforms (ODPs). Such popular systems are conceived in order to match individuals from two groups of potential mates. The business of such platforms pivots around the customers' expectancy to get in contact with their future dates: upon the payment of a fee to the platform owner, ODPs provide specific tools to improve reach and visibility**

However, ODPs require a critical number of active users in order to sustain their operations (and their business). Customers of the platform trade off on the price for being more visible and attract mates' contacts. A user becomes inactive if he or she is not contacted by others for some time: being contacted by potential mates acts as an activation signal.

The aim of our analysis is to propose a game theoretical framework to capture such a complex activation problem in strategic form. We unveil the structure of Nash equilibria and we further derive a Stackelberg formulation. The latter is a hierarchical game where the platform owner aims at maximizing profits while preserving the ODP activity level above a critical epidemic threshold.

***Index Terms*—Online Dating Platforms, Bipartite Graphs, Epidemics, Nash Equilibria, Stackelberg Equilibria.**

I. Introduction

The first online dating platforms (ODP) appeared during the 1990's and introduced online dating as an Internet service. Subscribers of such platforms access first a portal and create their own profile. By using the ODP, they seek for matches among the profiles of potential peer mates.

Beyond the over 5000 ODP platforms worldwide, the premier ODP websites in this domain are Match.com, OkCupid, and PlentyofFish. More recently, mobile dating apps have appeared to support the online dating market. Nowadays, ODP users can do everything from browsing profiles to setting up real-time dates conveniently from their smartphones. E.g., in April 2014, Match.com shipped the "Stream" location-based application.

According to recent surveys [1] *"Taken together, 11% of all American adults have used either an online dating site or a mobile dating app and are classified as online daters."*

In practice, the typical platform provides multiple service levels including various optional tools such as chatrooms, webcam calls or profile acceleration tools for ODP subscribers. All such tools are meant to let customers either increase their reach, e.g., by directly contacting potential mates, or becoming more visible in search results. Such optional tools require to pay a certain fee to the platform owner. Typically, the prototype ODP also provides information of potential matches by showing the tagged user with the username of the most recently registered community members, messages addressing the tagged user, the names of the most recent users who have browsed the tagged users profile. Also, such systems have recommenders' functionality to hint the names of mates which best match the tagged user; to this aim a notion of similarity is assessed through self-reported preferences over a set of predefined variables.

In Fig. 1 we have reported a typical dashboard presented to an ODP subscriber. We can identify there a *search* menu to perform search operations, an *email* menu to retrieve messages from peers through platform owned emails, a recommendation menu to check for best matches, and a tool to signal interest to peers (winks). The *boost* menu entry refers to an acceleration service able to give priority to certain subscribers' profiles and a *messenger* functionality provides a direct communication channel to peers.

From an abstract standpoint, an ODP could be truly defined as a matching engine. Applications of matching problems have been studied in science since early 1930s with the so called college admission problem [2]. Results in matching theory are often concerned with the existence of stable matchings and algorithms to obtain them.

In this paper, we are interested in a different problem: our question is rather when the activity level of ODP's subscribers is sufficient to sustain the operations of the platform. ODPs can function only if they have a sufficient number of *active* subscribers, i.e., those who take actions on the platform and contact potential partners. But, such activity has a cost. Chances to get in contact with potential mates increase with ODP's reach and visibility tools: in turn, those require the payment of a fee to the platform owner, e.g., enabling chats, messaging, etc.

Thus, the amount of active subscribers depends on the price they pay and on the attention they attain from their potential mates. Ultimately, the activity level of the platform connects to revenues generated for the platform owner.

Because of such mutual activation process, we model the ODP dynamics as an epidemic process in a bipartite graph corresponding to the two groups of potential mates. In this game, users of each group may decide how many of the platform's features use in order to increase their reach and visibility at a cost. We investigate the existence of equilibria

This work has been partially supported by the European Commission within the framework of the CONGAS project FP7-ICT-2011-8-317672, see www.congas-project.eu.

⋄CREATE-NET, via Alla Cascata 56c, 38100 Trento, Italy; ⋆INRIA Sophia-Antipolis, 2004 Route des Lucioles, 06902 Sophia-Antipolis Cedex, France; †EEMCS Delft University of Technology, Mekelweg 4 2628 CD Delft, The Netherlands.

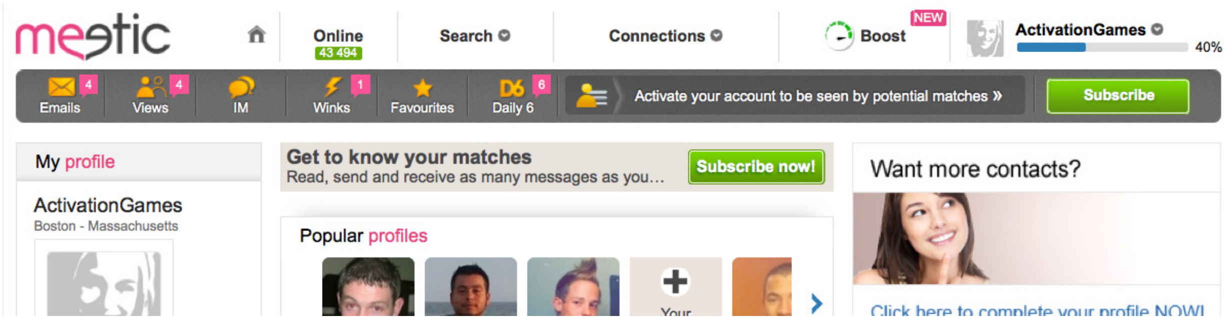

Figure 1: The subscriber dashboard of an online dating service platform. Observe in the uppermost part of the image the set of tools available to the customer: a search engine to explore profiles, a subscrbers' profile boost tool, platform-based emails, a tool to track own profile views and a messenger.

for the resulting game.

The paper's structure is as follows. In Sec. III we propose an epidemic spreading process on a bipartite graph to model the activation process dynamics. In Sec. IV we describe the game model and derive the structure of Nash equilibria. The strategic model is extended in Sec. V to account for strategic pricing at the platform owner side using a Stackelberg game. Finally, Sec. VII concludes the paper.

II. System Model

Potential mates of the ODP system form two groups of nodes, namely females and males, V_n, $n = 1, 2$ where $|V_n| = K_n$ is the number of subscribers of each group. For the sake of notation, we denote $\overline{n}$ as the opposite of group n, i.e., when $n = 1$ then $\overline{n} = 2$, and when $n = 2$ then $\overline{n} = 1$, respectively. Since contacts occur only between the two groups, we use a complete bipartite graph to represent the potential contacts between the two groups.

Each individual of the two groups of subscribers can be either active or inactive. A user is defined as active if he/she can perform actions such as browsing others' profile or contacting others. Moreover, we assume that only an active user may take the initiative to contact the others. The state of each individual i in group n is a binary random variable $X_n^{(i)}$ taking value 1 if the individual is active on the ODP or 0 if inactive.

An active user of group n becomes inactive after an exponentially distributed time with rate $\mu_n > 0$, $n = 1, 2$. An active user can activate an inactive user by browsing her/his profile and/or attempting to contact him/her (in ODPs this signal can be an email sent offline). We assume that, being contacted or browsed in profile, a member of group $\overline{n}$ is induced to perform some action on the ODP platform. Hence, he becomes active if he is inactive.

In order to model the action of an active user, we assume that an active individual explores the profiles of individuals in the opposite group according to a Poisson process. And, we let $\lambda_{n\overline{n}}^{(i)} > 0$ be the rate at which each active member i of group n makes contact to members of group $\overline{n}$, e.g., by browsing their profiles or making direct contact through the ODP platform.

Overall, this process is described by the probability of individual i in group n being active, $p_n^{(i)}(t) = P(X_n^{(i)}(t) = 1)$ which also implies that the probability of being inactive is $P(X_n^{(i)}(t) = 0) = 1 - p_n^{(i)}(t)$.

The underlying model is a continuous-time Markovian SIS model: as in the SIS model, a user can be repeatedly active, inactive and yet active again. Formally, the state of a node is thus described by a stochastic process of Bernoulli type $X_i(t)$. At time t we let $X_i(t) = 0$ if i is inactive and $X_i(t) = 1$ if i is active.

We further assume that activation and inactivation processes of all nodes are independent [3]. The rate a member of group n contacts/activates a member of group $\overline{n}$ is $\lambda_{n\overline{n}} > 0$ times per second.

As described in [3], [4], the SIS process developing on a graph with N nodes is actually a continuous-time Markov process with 2^N states. The dynamics of the nodal infection probability is derived by Kolmogorov differential equations. But, the resulting dynamical system consists of 2^N linear differential equations, not a viable solution for large networks. An alternative first order mean-field approximation is the NIMFA approach proposed by Van Mieghem et al. in [3], [4].

As a result, an activation signal is exchanged across the two groups in bipartite fashion: individuals of group n cannot activate individuals of the same group, but, they can activate individuals of the opposite group $\overline{n}$.

Individuals can increase the rate at which they become visible to the members of the opposite group. E.g., they can either use ODP functionalities such as messaging potential mates, they can browse more frequently becoming visible to their peers, or they can pay to make their profile becomes more visible through acceleration features. Doing so, they control the rate at which they are activated by members of the other group at a cost.[1]

Thus, $\lambda_{12}^{(i)}$ and $\lambda_{21}^{(i)}$ have the meaning of a continuous control variable which we denote *acceleration*. For practical reasons we assume that $0 \le \lambda_{21}^{(i)} \le \lambda_{21}^{\max}$ and $0 \le \lambda_{12}^{(i)} \le \lambda_{12}^{\max}$.

[1] We are referring to a monetary cost for the sake of simplicity. Observe that a more general setting could include possibly the budget of attention customers need to devote to be active on the ODP.

Depending on the amount of resources spent, each user will be able to reach more potential mates and become more visible: this generates a trade off between the aim of being contacted and the resources (time, money) spent in order to do so. Observe that the utility obtained by using a certain strategy will depend on the actions of members of the opposite group and of the same group as well. Actually, this provides a classic framework for the activation game proposed in the next section.

III. Game Model

The activation ODP game is defined as usual by players, strategies and utilities:

- *Players:* each member of the two groups of nodes V_n, $n = 1, 2$;
- *Strategies:* the rate $\lambda_{\overline{n}n}^{(h)}$ at which individual h of group n attract the attention of members of group $\overline{n}$;
- *Utility:* for an individual of group n it is an increasing function of the probability p_n of being activated minus the cost to obtain attention.

In the rest of the paper we assume that players have perfect information on the system's state and on the other players' strategies.

We can consider the general SIS model proposed in [3]:

$$\dot{p}_n^{(h)} = -\mu_n p_n^{(h)} + (1 - p_n n^{(h)})\lambda_{\overline{n}n}^{(h)} \sum_{k=1}^{K_{\overline{n}}} p_{\overline{n}}^{(k)}, \qquad (1)$$
$$n = 1, 2, \quad h = 1, \ldots, K_n$$

System (1) is a SIS type of epidemic equation developed within the NIMFA model framework. Such meanfield approximation corresponds to the Markovian dynamics introduced in the previous section. In our context, the original system requires a complete system of 2^N linear differential equations, $N = K_1 + K_2$. But, NIMFA represents a first order approximations of SIS dynamics: it requires that the infectious state of two nodes in the network are uncorrelated, namely $\mathrm{E}[X_i(t)X_j(t)] = \mathrm{E}[X_i(t)]\,\mathrm{E}[X_j(t)]$. Doing so, the original system is replaced by N non-linear differential equations (1).

Here, $\lambda_{n\overline{n}}^{(h)}$ is a constant, we can resort to the notion of metastable epidemic state. Implicitly, we are assuming that the time horizon is large enough to consider the metastable state of the resulting SIS system.

Actually, when the time horizon T is large enough, the meanfield approximation of the system (1) is driven into the quasi-stationary distribution [5] of the nodes' state, i.e., to the infection probability distribution denoted p_∞ in [3]. With some abuse of notation, in the rest of the paper, we will denote $p_\infty = p$ for notation's sake.

Overall, there are two possible solutions for the system (1): the all zero absorbing state and the non-zero quasi-stationary one when the system is above the epidemic threshold [3].

A. Homogeneous Nash Equilibria

For analytic tractability, in the rest of the paper we restrict our analysis to the case when individuals of the two groups adopt the same strategy and characterize equilibria under such assumption.

Once we dropped the user's index we can consider the following governing dynamics:

$$\begin{aligned} \dot{p}_1 &= -\mu_1 p_1 + (1 - p_1)\lambda_{21} K_2 p_2 \\ \dot{p}_2 &= -\mu_2 p_2 + (1 - p_2)\lambda_{12} K_1 p_1 \end{aligned} \qquad (2)$$

In particular, the system (2) has a non trivial solution only if Perron eigenvalue

$$\lambda_1\left(\begin{bmatrix} 0 & \rho_1 \\ \rho_2 & 0 \end{bmatrix}\right) > 1$$

where $\rho_n := K_{\overline{n}}\lambda_{\overline{n}n}/\mu_n$, equivalently if and only if $\lambda_{21} > 0$: in particular, if and only if $\rho_1\rho_2 > 1$. For notation's sake, $\rho^{\max} := K_{\overline{n}}\lambda_{\overline{n}n}^{\max}/\mu_n$. Otherwise, the trivial zero solution is the unique asymptotically stable restpoint for the original system.

The condition can hence be resumed as

Proposition 1 (Epidemic Threshold)**.** *The system* (1) *has a non trivial quasi-stationary probability distribution if and only if*

$$\rho_1\rho_2 > 1 \qquad (3)$$

In order to avoid degenerate cases, in the rest of our discussion, we need to assume that $\rho_1^{\max}\rho_2^{\max} > 1$. Above the threshold, the unique solution of (2) writes

$$p_n = \frac{\rho_n\rho_{\overline{n}} - 1}{\rho_{\overline{n}}(1 + \rho_n)}, \; n = 1, 2 \qquad (4)$$

from which the utility

$$\tilde{S}_n = \Big(\frac{\rho_n\rho_{\overline{n}} - 1}{\rho_{\overline{n}}(1 + \rho_n)}\Big) - \gamma_n\rho_n \qquad (5)$$

where the linear cost term $\gamma_n := \delta_n\mu_n/K_{\overline{n}}$ is including the normalization term for notation's sake. Using standard game theoretical settings, we assume that each individual of the two groups acts in order to maximize their own utility.

The actual utility writes

$$S_n = \left[\Big(\frac{\rho_n\rho_{\overline{n}} - 1}{\rho_{\overline{n}}(1 + \rho_n)}\Big) - \gamma_n\rho_n\right]^+ \qquad (6)$$

where $[x]^+ = \max\{0, x\}$, because $\rho_n = 0$ leads to the die out of the epidemic according to (3), resulting in utility zero.

In order to proceed further we need to find the maximum of (5), and then apply the results to find the best response and Nash equilibria, with some care for the case when $\tilde{S}_n < 0$. Actually, one can differentiate (5) for a fixed strategy of player $\overline{n}$, and obtain the best response of player n. The best response is determined by maximizing the utility of a player for a fixed opponent's strategy $\rho_{\overline{n}}$

$$\widehat{\rho}_n = -1 + \sqrt{\frac{1 + \rho_{\overline{n}}}{\gamma_n\rho_{\overline{n}}}} \qquad (7)$$

Now, a direct calculation shows that

$$\frac{d^2\tilde{S}_n}{d\rho_n^2} = -2\frac{1 + \rho_{\overline{n}}}{(1 + \rho_n)^3\rho_{\overline{n}}} < 0$$

Hence, $\widehat{\rho}_n$ maximizes (5). However, the limit of $\rho_n \leq \rho_n^{\max}$ and whether the condition for the epidemic to spread out are both ignored.

In order to perform the analysis of the possible Nash equilibria, we need to determine the best response ρ^* of player n to the opponent's strategy $\rho_{\overline{n}}$. We report here the simplest form of the best response result, which provides the results useful for the following derivations; a more detailed description is reported in the Appendix.

Taking into account the constraint for strategy $0 \leq \rho_n \leq \rho_n^{\max}$, and the epidemic threshold condition (3)

Proposition 2. *The best response of subscribers of group n:*
i. $\rho_n^* = 0$ if $\rho_{\overline{n}} = 0$ or if $\gamma_n \geq 1$ and $\rho_{\overline{n}} \geq (\gamma_n - 1)^{-1}$.
ii. $\rho_n^* = \widehat{\rho}_n$, if and only if $\rho_{\overline{n}}^{-1} < \gamma_n(1 + \rho_n^{max}) - 1$ and $\widehat{\rho}_n \rho_{\overline{n}} > 1$
iii. $\rho_n^* = \rho_n^{max}$, if and only if $\rho_{\overline{n}}^{-1} \geq \gamma_n(1 + \rho_n^{max}) - 1$ and $\rho_n^{max} \rho_{\overline{n}} > 1$

Proof. i. If $\rho_{\overline{n}} = 0$ we already found that $\rho_n^* = 0$. Hence, let us assume that $\rho_{\overline{n}} > 0$: the condition $\rho_n^* = 0$ is equivalent to state $\widehat{\rho} \leq 0$. But, for $\gamma_n < 1$, it follows $1 + \rho_{\overline{n}}^{-1} \leq \gamma_n$ which leads to a contradiction with the fact that $\rho_{\overline{n}} > 0$. Conversely, if $\gamma_n \geq 1$, it must hold $\rho_{\overline{n}} > 1/(\gamma_n - 1)$.
ii. Follows directly from the condition $\widehat{\rho} \leq \rho_n^{\max}$, whereas iii. is the complementary case of ii.
Finally, we need to enforce (3) into conditions ii. and iii. in order to let the system above the epidemic threshold. □

In the next section we analyze the Nash equilibria resulting from the game.

IV. Nash Equilibria

The Nash equilibria of the game can be resumed by the following results where we distinguish between *interior* and *extremal* Nash equilibria. In particular, the first result is about *extremal* Nash equilibria, i.e., those where one of the players adopts either $\rho_n = 0$ or $\rho_n = \rho_n^{\max}$. Interior Nash equilibria are those which are not extremal.

Theorem 1 (Extremal Nash Equilibria)**.** *Extremal Nash equilibria are as follows:*
Type i. $\mathbf{s} = (0, 0)$ is always a Nash equilibrium.
Type ii. If $\gamma_1, \gamma_2 < 1$, an extremal Nash equilibrium exists in the form

$$\rho_n^* = \rho_n^{max}, \quad n = 1, 2$$

if and only if

$$\frac{1}{\gamma_n} \geq \frac{(1 + \rho_n^{max})^2 \rho_{\overline{n}}^{max}}{1 + \rho_{\overline{n}}^{max}}, \;\; n = 1, 2$$

If $\gamma_1, \gamma_2 \geq 1$, no such Nash equilibrium exists.
Type iii. A Nash equilibrium in the form

$$\rho_n^* = -1 + \sqrt{\frac{1 + \rho_{\overline{n}}^{max}}{\gamma_n \rho_{\overline{n}}^{max}}}, \quad \rho_{\overline{n}}^* = \rho_{\overline{n}}^{max}$$

exists iff $\gamma_n, \gamma_{\overline{n}} < 1$ and

$$\frac{1 + \rho_{\overline{n}}^{max}}{\rho_{\overline{n}}^{max}} < \frac{1}{\gamma_n} < \frac{(1 + \rho_n^{max})^2 \rho_{\overline{n}}^{max}}{1 + \rho_{\overline{n}}^{max}}$$

and

$$\frac{1}{\gamma_{\overline{n}}} \geq (1 + \rho_n^{max})^2 \frac{\rho_n^*}{1 + \rho_n^*}$$

If $\gamma_n < 1$ and $\gamma_{\overline{n}} \geq 1$, the additional condition is required

$$\frac{\gamma_{\overline{n}} - 1}{\gamma_n \gamma_{\overline{n}}} < \frac{\rho_{\overline{n}}^{max}}{1 + \rho_{\overline{n}}^{max}}$$

If $\gamma_n \geq 1$, no such Nash equilibrium exists.

Proof. Type i. It is easy to verify that zero is always a Nash equilibrium: in fact, any unilateral deviation does not increase p_n, but incurs a positive cost.
Type ii. It requires a direct verification for the condition $\widehat{\rho}_n > \rho_n^{\max}$ to hold; in the case when $\gamma_n > 1$, $n = 1, 2$, the negative result comes from Prop. 2, case i.
Type iii. Similar to the previous two cases: the last two conditions require again to impose the conditions implied by Prop. 2, case i. We further observe that

$$\frac{1 + \rho_{\overline{n}}^{\max}}{\rho_{\overline{n}}^{\max}} < \frac{(1 + \rho_n^{\max})^2 \rho_{\overline{n}}^{\max}}{1 + \rho_{\overline{n}}^{\max}}$$

is equivalent to the condition $\rho_1^{\max} \rho_2^{\max} > 1$. □

Clearly, because the best response to $\rho_{\overline{n}} = 0$ is $\rho_n = 0$, no equilibrium of the type $\rho_n = 0$ and $\rho_{\overline{n}} > 0$, $n = 1, 2$ is possible.

However, interior Nash equilibria cannot always be derived for (2).

Theorem 2 (Interior Nash Equilibria)**.** *The following two cases hold:*
i. If $\gamma_i > 1$ for either $i = 1$ or $i = 2$ or both, no interior Nash equilibrium is possible.
ii. If $\gamma_i \leq 1$ for $i = 1, 2$,

$$\frac{1 + \rho_{\overline{n}}^{max}}{\rho_{\overline{n}}^{max}} < \frac{1}{\gamma_n} < \frac{(1 + \rho_n^{max})^2 \rho_{\overline{n}}^{max}}{1 + \rho_{\overline{n}}^{max}} \tag{8}$$

at most a unique interior Nash equilibrium exists.

Proof. i. By writing the condition for the interior Nash equilibrium $0 < \rho_n < \rho_n^{\max}$, $n = 1, 2$, we obtain simply

$$0 \leq \frac{1}{\gamma_{\overline{n}}(1 + \rho_{\overline{n}}^{\max})^2 - 1} < \rho_n < \frac{1}{1 - \gamma_{\overline{n}}}, \; n = 1, 2$$

The above relation proves that no Nash equilibrium is possible if $\gamma_n > 1$ for either $n = 1$ or $n = 2$.
ii. The bounds in (8) force an interior Nash equilibrium. We further observe that an interior Nash must solve the following system

$$\gamma_n (1 + \rho_n)^2 = \left(1 - \sqrt{\frac{\gamma_{\overline{n}} \rho_n}{1 + \rho_n}}\right)^{-1}, \quad n = 1, 2$$

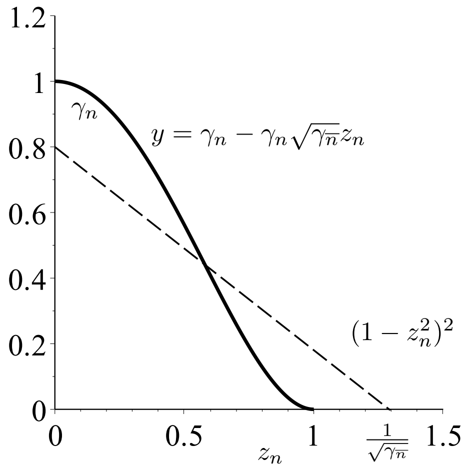

Figure 2: Interior Nash Equilibria: a unique interior Nash is possible when $\gamma_n \le 1$, $z_n \ge \sqrt{\gamma_{\overline{n}}}$, $n = 1, 2$.

under the constraint $0 < \rho_n < \rho_n^{\max}$. The same system can be conveniently rewritten by replacing $z_n := \sqrt{\frac{\rho_n}{\rho_n+1}}$, where we hence obtain the equivalent one

$$\gamma_n(1 - \sqrt{\gamma_{\overline{n}}}z_n) = (1 - z_n^2)^2 \tag{9}$$

Solutions z_n of (9) exists if $f(z_n) = (1 - z_n^2)^2$ and the parametric line $y := \gamma_n - \gamma_n\sqrt{\gamma_{\overline{n}}}z_n$ intersect for $0 < z_n < 1$. It can be showed that the solution of (9) always exists: it actually corresponds to graphically, as depicted in Fig. 2. The fact that it is unique is not proved here for the sake of space but it can be derived by simple arguments on the derivatives of the curves depicted in Fig. 2.

In order to exist, the interior Nash equilibrium should also satisfy (3), which writes $\rho_n\Big(-1+\sqrt{\frac{1+\rho_n^{-1}}{\gamma_{\overline{n}}}}\Big) > 1$ in this case. From such condition, it is immediate to derive the condition $\rho_n > \frac{\gamma_{\overline{n}}}{1-\gamma_{\overline{n}}}$, $n = 1, 2$, which also writes $z_n > \sqrt{\gamma_{\overline{n}}}$, $n = 1, 2$. □

Remark 1. *We can resume the above results saying that the ODP activation game has at most two Nash equilibria. It always has at least one Nash equilibrium, i.e., the null Nash equilibrium. In addition, a non null Nash equilibrium is possible depending on the range of parameters γ_n, $n = 1, 2$. Furthermore, such equilibrium is extremal and unique when either one of $\gamma_n \ge 1$. When $\gamma_n < 1$ for $n = 1, 2$, the equilibrium is an interior one.*

V. Stackelberg Equilibrium

We now assume the ODP owner acts such in a way to generate the largest possible revenue from active users. In order to do so, the platform owner can act on the price subscribers pay.

The pricing determines the revenue generated by charging subscribers who make use of optional system features, i.e., the tools made available in order to enhance users' reach and visibility. E.g., acceleration features are customary in ODPs, however a cost has to be paid in order to appear higher in the recommendation list seen by potential mates. The strategic interaction between the platform owner and the two groups of subscribers is a hierarchical game where actions occur sequentially: first prices are decided by the platform owner, and then subscribers decide how much they make use of optional system features.

The natural framework to model this strategic interaction is the notion of a Stackelberg game [6], where the platform owner, i.e., the leader, makes the first move. In this framework, it settles (γ_1, γ_2), i.e., the cost paid by a customer of each group in order become more visible using the acceleration tools.

The two groups react, as followers, and attain a certain Nash equilibrium. The ODP owner utility is thus determined by the pair $\rho_n(\gamma_n, \gamma_{\overline{n}})$, $n = 1, 2$, i.e., the followers' response to the leader setting the service platform costs.

Hence, we study the Stackelberg equilibrium where the broker aims at solving the following problem

$$\max_{\gamma_1,\gamma_2} S_p(\gamma_1, \gamma_2) = \gamma_1\rho_1(\gamma_1, \gamma_2) + \gamma_2\rho_2(\gamma_1, \gamma_2) \tag{10}$$

where maximizing (10) corresponds to maximizing the profit of the platform owner.

In order to solve the Stackelberg equilibrium, we have to determine the optimal pair: to do so we analyze the utility based on the type of Nash equilibrium which is induced by the leader's move.

Remark 2. *Typically, a well-posed Stackelberg problem* (10) *requires a unique Nash to be induced. In our case, because the non-null Nash equilibrium is always a solution, and at most two Nash equilibria exist, we should optimize versus the largest induced Nash. Actually, from the results in Thm. 2 and Thm. 1, we know that a bijection exists between the pairs (γ_1, γ_2) and the corresponding non-zero Nash equilibria. This is the key property which enables our further analysis of the Stackelberg equilibria.*

We start discussing the possible cases in the following, enumerating the possible induced Nash equilibria.

Type i. Cost pairs inducing the null Nash equilibrium only, i.e., $\rho_n = 0$, $n = 1, 2$, correspond to $S_p = 0$, and cannot be optimal. Actually, the leader can always find a strategy that is better off: as seen in the following case (γ_1, γ_2) can be always chosen in order to induce a Nash equilibrium with positive utility.

Type ii. For the case of equilibria of type two, the utility of the platform owner writes

$$S_p(\gamma_1, \gamma_2) = \gamma_1\,\rho_1^{\max} + \gamma_2\,\rho_2^{\max} \tag{11}$$

under the constraints $\gamma_n \le \frac{1+\rho_{\overline{n}}^{\max}}{\rho_{\overline{n}}^{\max}}(1 + \rho_n^{\max})^{-2}$, $n = 1, 2$. Observe that $(1 + \rho_n^{\max})^2 > (1 + \rho_n^{\max}) > (1 + 1/\rho_{\overline{n}}^{\max})$ because $\rho_1^{\max}\rho_2^{\max} > 1$. Hence, the maximum is

$$S_p(\gamma_1, \gamma_2) = \sum_{n=1,2} \frac{1 + \rho_{\overline{n}}^{\max}}{\rho_{\overline{n}}^{\max}} \frac{\rho_n^{\max}}{(1 + \rho_n^{\max})^2} \tag{12}$$

which corresponds to the platform owner utility for the cost values attaining equality in the constraint.

Type iii. We consider equilibria of the type $\widehat{\rho}_1, \rho_2^{\max}$ for the sake of simplicity. Here, the leader chooses the cost pair in the set where the constraints in Thm. 1 hold. The form of the utility is in this case:

$$S_p(\gamma_1, \gamma_2) = \gamma_1 \left(-1 + \sqrt{\frac{1 + \rho_2^{\max}}{\gamma_1 \rho_2^{\max}}} \right) + \gamma_2\, \rho_2^{\max} \tag{13}$$

By replacing the expression for $\widehat{\rho}_1 = -1 + \sqrt{\frac{1+\rho_2^{\max}}{\gamma_1 \rho_2^{\max}}}$, and maximizing for γ_2 as from the constraint in Thm.1 we can recognize that the maximum of S_b is indeed attained for

$$\gamma_2 = \frac{1}{(1 + \rho_1^{\max})^2} \Big(1 - \sqrt{\frac{\gamma_1 \rho_2^{\max}}{1 + \rho_2^{\max}}} \Big)$$

Note that based on Thm.1, we need to assume $\gamma_1 < 1$, so that $\gamma_2 < 1$ as well from the previous expression. This rules out the case $\gamma_2 \geq 1$.

We can thus derive the expression for the utility as a function of the cost of group V_1 only, namely,

$$S_p(\gamma_1, \gamma_2) = \quad -\gamma_1 + \sqrt{\gamma}_1 \left[\sqrt{\frac{1+\rho_2^{\max}}{\rho_2^{\max}}} - \frac{\rho_2^{\max}}{(1+\rho_1^{\max})^2} \sqrt{\frac{\rho_2^{\max}}{1+\rho_2^{\max}}} \right] + \frac{\rho_2^{\max}}{(1+\rho_1^{\max})^2}$$

It is possible to see that the absolute maximum of such function is attained at

$$\widehat{\gamma_1} = \frac{1}{4} \left[\sqrt{\frac{1 + \rho_2^{\max}}{\rho_2^{\max}}} - \frac{\rho_2^{\max}}{(1 + \rho_1^{\max})^2} \sqrt{\frac{\rho_2^{\max}}{1 + \rho_2^{\max}}} \right]^2$$

Hence, based on the constraints described in Thm. 1 there are three cases determined by the relative position of $\widehat{\gamma_1}$ with respect to the interval $[\widetilde{\gamma_1}\ \overline{\gamma_1}]$ where

$$\widetilde{\gamma_1} := \frac{1 + \rho_{\overline{n}}^{\max}}{(1 + \rho_n^{\max})^2 \rho_{\overline{n}}^{\max}} \qquad \overline{\gamma_1} := \frac{\rho_{\overline{n}}^{\max}}{1 + \rho_{\overline{n}}^{\max}}$$

- If $\widetilde{\gamma_1} < \widehat{\gamma_1} < \widehat{\gamma_1}$, then $\gamma_1^* = \widehat{\gamma_1}$ and the optimal leader's utility is given by (14);
- if $\widehat{\gamma_1} < \widetilde{\gamma_1}$, the leader can make the utility arbitrarily close to $\tilde{S}_p(\rho_1^{\max}, \rho_2^{\max})$ by setting $\gamma_1 = \widetilde{\gamma_1} + \epsilon$, with $\epsilon > 0$ a fixed arbitrarily small value. But, then the induced Nash equilibrium is of Type ii.
- if $\widehat{\gamma_1} > \overline{\gamma_1}$, the leader can make the the utility arbitrarily close to $\tilde{S}_p(\overline{\gamma_1}, \rho_2^{\max})$ by setting $\gamma_1 = \overline{\gamma_1} - \epsilon$, with $\epsilon > 0$ a fixed arbitrarily small value. But, we recall that the bound $\overline{\gamma_1}$ corresponds to the condition $\inf_{\rho_1}\{\rho_1 \rho_2^{\max}\} > 1$. Induced Nash equilibrium of Type ii would indeed be better off, so that this case cannot correspond to a Stackelberg equilibrium.

We can finally map the set of pairs (γ_1, γ_2) onto the most profitable extremal Nash equilibrium as a function of $(\rho_1^{\max}, \rho_2^{\max})$:

$$\phi(\gamma_1, \gamma_2) = \begin{cases} (\widehat{\rho}_1, \rho_2^{\max}) & \text{if} \quad \bar{\gamma}_1 < \gamma_1 < \widetilde{\gamma}_1, \gamma_2 = \widetilde{\gamma}_2 \\ (\rho_1^{\max}, \widehat{\rho}_2) & \text{if} \quad \gamma_1 = \widetilde{\gamma}_1, \bar{\gamma}_2 < \gamma_2 < \widetilde{\gamma}_2 \\ (\rho_1^{\max}, \rho_2^{\max}) & \text{otherwise} \end{cases} \tag{14}$$

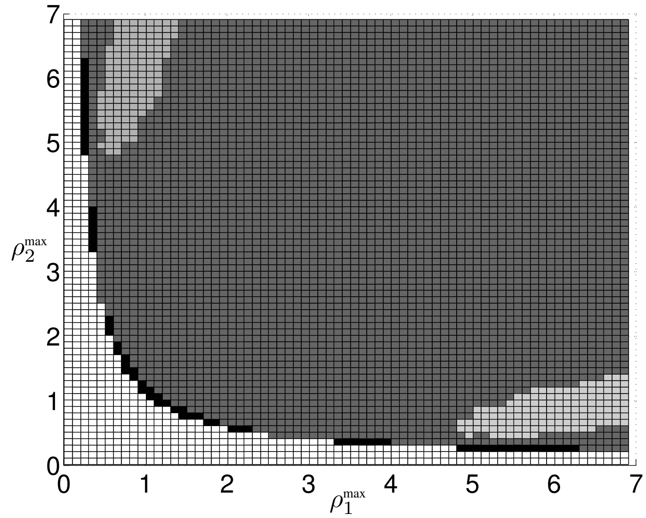

Figure 3: Stackelberg-induced Nash equilibria as a function of $\rho_1^{\max}$ and $\rho_2^{\max}$: the black area represents type ii induced Nash equilibria $(\rho_1^{\max}, \rho_2^{\max})$, the light gray area represents type iii induced equilibria of the kind $(\rho_1^{\max}, \widehat{\rho}_2)$ and of the kind $(\widehat{\rho}_1, \rho_2^{\max})$; the dark gray area represents the interior Nash equilibria.

Finally, we have reported in Fig. 3 the map of the type of induced non zero Nash equilibria depending on $\rho_i^{\max}$, $i = 1, 2$.

VI. Related Works

With respect to online dating platforms, most attention in literature has been attracted by sociological and psychological issues [7], [8]. Despite those are out of the scope of our analysis, such considerations are indeed relevant to design ODP's matching mechanisms. Furthermore, users' preferences and profiles are crucial for the data mining process used in order to feed matching algorithms in online dating platforms. Conversely, a vast literature in computer science covers matching problems and design of matching algorithms starting from mutual interest of agents. The standard prototype of a matching algorithm is described in [2].

Works related to ours are [9], [10]. In [9] the focus is on the match-making algorithms based on machine learning techniques. Many features of ODPs are described in the seminal work [10] which reports on an early Internet dating community in Sweden. The authors perform a complex network analysis unveiling the typical topological traits derived from the relations within such platforms.

Diffusion of epidemic processes on graphs attracted the attention of the scientific community; the NIMFA model is described in [3][4]. The general problem of approximation of a epidemic process via ODEs is discussed in several works, see for instance [11].

Processes developing over bipartite graphs have been studied in literature before. The analysis of popularity for content diffused on bipartite networks has been carried out in [12]. The authors refer to the NetFlix's content distribution platform, where videos and users represent the partition of the graph.

To the best of the authors' knowledge, this is the first work to provide a model based on the strategic interaction of actors in online dating platforms.

VII. Conclusions

Our initial study of online dating platforms provides a novel perspective on the relation between the number of users and pricing of ODP's services. We have factored into our model the strategic interaction of ODP users, who try to attract the attention of potential mates at a cost settled by the platform owner. We have showed that reciprocal activation of the two groups introduces a SIS type of dynamics. Such dynamics settles a lower bound in the activation of platform's customers below which the system switches off.

Actually, because active users represent the source of revenue for ODPs, maintaining a large set of active subscribers is the main target of the platform owner. This work suggests several interesting research directions for the design and the pricing models for ODPs. The first immediate extension is including the presence of non group-homogeneous Nash equilibria for the game. Moreover, we have been studying the system under the customary assumption that players are rational and have perfect information on the structure of the game. Mechanism design and signaling games appear hence another extension that is worth exploring for the activation game.

References

[1] A. Smith and M. Duggan, "Online dating & relationships," Pew Research Center, Tech. Rep., 21 October 2013.

[2] D. Gale and L.S. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, no. 69, p. 915, 1962.

[3] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 1, pp. 1–14, Feb 2009.

[4] P. Van Mieghem, "The n-intertwined sis epidemic network model," *Computing*, vol. 93, no. 2-4, pp. 147–169, 2011. [Online]. Available: http://dx.doi.org/10.1007/s00607-011-0155-y

[5] J. N. Darroch and E. Seneta, "On quasi-stationary distributions in absorbing discrete-time finite Markov chains," vol. 2, no. 1, pp. 88–100, Jun. 1965. [Online]. Available: http://www.jstor.org/stable/3211876

[6] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1999.

[7] C. L. Toma, J. T. Hancock, and N. B. Ellison, "Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles," *Personality and Social Psychology Bulletin*, vol. 34, no. 8, pp. 1023–1036, 2008.

[8] M. J. Rosenfeld and R. J. Thomas, "Searching for a mate the rise of the internet as a social intermediary," *American Sociological Review*, vol. 77, no. 4, pp. 523–547, 2012.

[9] F. Diaz, D. Metzler, and S. Amer-Yahia, "Relevance and ranking in online dating systems," in *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '10. New York, NY, USA: ACM, 2010, pp. 66–73. [Online]. Available: http://doi.acm.org/10.1145/1835449.1835463

[10] P. Holme, C. R. Edling, and F. Liljeros, "Structure and time evolution of an internet dating community," *Social Networks*, vol. 26, no. 2, pp. 155 – 174, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378873304000085

[11] A. Mandelbaum and G. Pats, "State-dependent stochastic networks. part i. approximations and applications with continuous diffusion limits," *Ann. Appl. Probab.*, vol. 8, no. 2, pp. 569–646, 05 1998. [Online]. Available: http://dx.doi.org/10.1214/aoap/1028903539

[12] M. Beguerisse Diaz, M. A. Porter, and J. Onnela, "Competition for popularity in bipartite networks," *Chaos*, vol. 20, no. 043101, 2010.

Appendix

Proposition 3. *The best response of the subscribers of group n to the strategy $\rho_{\overline{n}}$ of the other group $\overline{n}$ is:*

i. $\rho_n^* = \widehat{\rho_n}$, *if and only if* $\gamma_n < 1$ *and* $\rho_{\overline{n}} > \max\left(\frac{1}{\gamma_n(1+\rho_n^{\max})^2-1}, \frac{4}{\frac{1}{\gamma_n}+\gamma_n-2}\right)$;

ii. $\rho_n^* = \rho_n^{\max}$, *if and only if* $\gamma_n < 1$, $\frac{(1-\gamma_n)\rho_{\overline{n}}-\sqrt{(\gamma_n-1)^2\rho_{\overline{n}}^2-4\gamma_n\rho_{\overline{n}}}}{2\gamma_n\rho_{\overline{n}}} < \rho_{\overline{n}} < \frac{(1-\gamma_n)\rho_{\overline{n}}+\sqrt{(\gamma_n-1)^2\rho_{\overline{n}}^2-4\gamma_n\rho_{\overline{n}}}}{2\gamma_n\rho_{\overline{n}}}$ *and* $\rho_{\overline{n}} \le \max\frac{1}{\gamma_n(1+\rho_n^{\max})^2-1}$;

iii. otherwise, $\rho_n^* = 0$.

Proof. When $\gamma_n \ge 1$, we have $\widetilde{S_n} < 0$ for any positive ρ_n. Hence, the best response is $\rho_n^* = 0$ when $\gamma_n \ge 1$. If $\rho_{\overline{n}} = 0$ we already found that $\rho_n^* = 0$. We focus thus on the case when $\gamma_n < 1$ and $\rho_{\overline{n}} > 0$ to explore non zero best response. The best response is neither 0 nor $\rho_n^{\max}$ but $\widehat{\rho_n}$, i.e. $0 \ne \rho_n^* = \widehat{\rho_n} \ne \rho_n^{\max}$ if and only if $0 < \widehat{\rho_n} < \rho_n^{\max}$, $\widetilde{S_n} > 0$ and $\rho_n\rho_{\overline{n}} > 1$ when $\rho_n = \widehat{\rho_n}$. When $\gamma_n < 1$ and $\rho_{\overline{n}} > 0$, $\widehat{\rho_n} > 0$ according to (7). The condition $\widehat{\rho_n} < \rho_n^{\max}$ leads to $\rho_{\overline{n}} > \frac{1}{\gamma_n(1+\rho_n^{\max})^2-1}$ based on (7) and $\widetilde{S_n} > 0$ when $\rho_n = \widehat{\rho_n}$ requires $\rho_{\overline{n}} > \frac{4}{\frac{1}{\gamma_n}+\gamma_n-2} > 0$ by combining (7) and (5). When $\rho_n = \widehat{\rho_n} > 0$ and $\widetilde{S_n} > 0$, the condition that $\rho_n\rho_{\overline{n}} > 1$ i.e. the system should be above the epidemic threshold is naturally satisfied according to (5). Similarly, $\rho_n^* = \rho_n^{\max}$ if and only if $\widetilde{S_n} > 0$ when $\rho_n = \rho_n^{\max}$ and $\widehat{\rho_n} \ge \rho_n^{\max}$.The former requires $\frac{(\gamma_n-1)\rho_{\overline{n}}-\sqrt{(\gamma_n-1)^2\rho_{\overline{n}}^2-4\gamma_n\rho_{\overline{n}}}}{-2\gamma_n\rho_{\overline{n}}} < \rho_n^{\max} < \frac{(\gamma_n-1)\rho_{\overline{n}}+\sqrt{(\gamma_n-1)^2\rho_{\overline{n}}^2-4\gamma_n\rho_{\overline{n}}}}{-2\gamma_n\rho_{\overline{n}}}$. The latter leads to $\rho_{\overline{n}} \le \frac{1}{\gamma_n(1+\rho_n^{\max})^2-1}$. $\square$