Steering User Behavior with Badges*

Ashton Anderson
Stanford University
ashton@cs.stanford.edu

Daniel Huttenlocher
Cornell University

Jon Kleinberg
Cornell University

{dph, kleinber}@cs.cornell.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

ABSTRACT

An increasingly common feature of online communities and social media sites is a mechanism for rewarding user achievements based on a system of *badges*. Badges are given to users for particular contributions to a site, such as performing a certain number of actions of a given type. They have been employed in many domains, including news sites like the Huffington Post, educational sites like Khan Academy, and knowledge-creation sites like Wikipedia and Stack Overflow. At the most basic level, badges serve as a summary of a user's key accomplishments; however, experience with these sites also shows that users will put in non-trivial amounts of work to achieve particular badges, and as such, badges can act as powerful incentives. Thus far, however, the incentive structures created by badges have not been well understood, making it difficult to deploy badges with an eye toward the incentives they are likely to create.

In this paper, we study how badges can influence and steer user behavior on a site—leading both to increased participation and to changes in the mix of activities a user pursues on the site. We introduce a formal model for reasoning about user behavior in the presence of badges, and in particular for analyzing the ways in which badges can steer users to change their behavior. To evaluate the main predictions of our model, we study the use of badges and their effects on the widely used Stack Overflow question-answering site, and find evidence that their badges steer behavior in ways closely consistent with the predictions of our model. Finally, we investigate the problem of how to optimally place badges in order to induce particular user behaviors. Several robust design principles emerge from our framework that could potentially aid in the design of incentives for a broad range of sites.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics.

General Terms: Algorithms, Economics, Theory.

Keywords: Badges, steering, badge placement problem.

1. INTRODUCTION

Give me enough medals and I'll win you any war.
—Napoleon

*This is an extended version of what appears in the conference proceedings, the only difference being the inclusion of two appendices that spell out details we had to omit in the conference version due to space constraints.

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2013, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2035-1/13/05.

Designers of online communities and social media sites have increasingly been making use of *badges* as a way to reward users for their achievements. Badges have been employed across a wide range of domains, from news sites like Huffington Post, where users are recognized for contributing valued comments and being well-connected; to education sites like Khan Academy, where users are awarded badges for watching instructional videos and correctly answering questions; to knowledge creation sites like Wikipedia and Stack Overflow, where users are awarded for their contributions to the online community. The use of badges in these settings invokes a much longer history of badge use in off-line domains: for example, organizations like the Boy Scouts award (literal) badges for proficiency at particular tasks, and airlines award elite status for specific amounts of travel.

Badges play multiple roles in all these settings. First, they function as a credentialing system, summarizing the skills and achievements of the individuals who receive them. But they also work powerfully as *incentives*; experience across many domains shows that people will direct considerable amounts of effort in pursuit of a badge. It is this incentive function of badges, and particularly the ways these badge-based incentives can be used in online applications, that is the focus of our work here—the ways in which badges can affect the extent and form of user participation in online communities and social media sites.

Badges as Incentives. The question of user participation and contribution in online domains is a broad topic that a number of active lines of research have contributed to, including social-psychological studies of user engagement [5, 6], incentives for effort and high-quality contribution in social media [12], mechanisms for distributed online recruitment [2, 7, 13], and the design of contests to encourage crowdsourced effort [8]. Different systems have been used successfully in different settings.

Badges are in several respects simpler than some of these other incentive mechanisms, lacking the direct competition of auctions and leaderboards and the exchangeability of currency-like systems. Despite their simplicity, however, in practice many social sites have positioned badges as an important part of their incentive systems. Furthermore, they appear to induce complex user behaviors that are thus far not well understood. Understanding this complexity, and developing principles for reasoning about how to use badges most effectively, are key underlying motivations of the present research.

Perhaps the most basic way badges can affect user behavior is by encouraging users to increase their overall level of participation. But much of the richness and complexity of badges comes from the fact that they are intrinsically based on a kind of multi-faceted way of thinking, in the following sense. A site will generally have many different types of activities that users can perform (the different dimensions of possible contributions to the site), and by creating badges that reward certain of these contributions at specific levels, the site's designers can attempt to "steer" a user's activities toward particular forms of contribution. For example, a question-answering (Q&A) site may have a range of activities that include asking questions, answering questions, up-voting and down-voting questions and answers, and others. On a Q&A site where everyone wants to ask questions and few people want to answer them, introducing a badge for users who have contributed a certain number of answers can steer the community toward contributing to this underrepresented action type. Badges for voting can similarly try to steer users toward providing enough feedback to maintain a useful quality signal on the content.

In summary, badges provide a rich language for expressing incentives, but with little existing framework for reasoning about their effects. Our work addresses a set of questions that can help provide insight into badges and their use. In particular, a natural first question is: Do badges work? That is, can we find concrete evidence that badges increase site participation or steer users towards taking actions they might not have taken otherwise? If badges do have an effect on users, how can we model user behavior in the presence of badges? And to the extent that designers can indeed steer user behavior with badges, how should they define badges to achieve the outcomes they want?

Overview of Results

Our work is based on this set of questions, and comprises three different components. First, we develop a model for user behavior in the presence of badges on a site with multiple types of activities. Second, we evaluate our model on data from the popular question-answering site Stack Overflow, and show that the main qualitative predictions of the model match what we observe in the aggregate user behavior on the real site. Third, we consider how a site designer can use such a model to define badges with the goal of achieving a desired pattern of behavior.

Theoretical Model of User Behavior. Given the discussion of badges thus far, what are the basic ingredients we need in order to define a model of user behavior with badges? Intuitively, we would like to have a multi-dimensional space representing the possible types of actions on the site; users have a preferred mixture of activities in this space, and introducing badges can induce them to shift their mixture in particular ways. Our model brings these features together in a natural way.

Since a key aspect of badges is the way in which different badges reward different activities, we model the site as having n different *action types*; as in our example of a Q&A site above, the action types could correspond to asking a question, answering a question, voting on a question or answer, and others. Users perform actions, choosing from among these types; a user's mix of actions over his lifetime can thus be thought of as defining a vector in an n-dimensional space whose i^{th} coordinate records the number of actions of type i that he has taken.[1] Each badge is defined by specifying how many actions of each type must be performed in order for it to be awarded; in this sense, each badge defines a "frontier" in the space that the user's vector must cross in order for the badge to be awarded. We refer to this frontier as the *badge boundary*.

In order to talk about the incentives that badges create, we also need a model of a user's utility, which we represent in two components. First, each user has a preferred mix of actions on the site, corresponding to a distribution over the possible action types; in the absence of badges, he would simply perform actions according to this mixture by sampling from his distribution over possible action types. Sampling action types from a different distribution incurs a cost to the user, based on how far the distribution is from his preferred one. Second, a badge confers utility once it is awarded, and so a user has an incentive to shift his distribution of activities somewhat in order to achieve the badge more quickly, which trades off against the cost for deviating from the preferred distribution.

The user's behavior is thus determined by the solution to an individual-level optimization problem, trading off the cost of shifting his activity distribution against the value of the badge. We will see that when this optimization is solved, the force of the incentives from badges clearly emerges naturally: the user's distribution of actions is deflected in the direction of badges, and there is an "acceleration" effect in which this deflection becomes stronger as the user approaches a badge boundary.

Thus far the model captures how a badge can steer users toward certain types of actions. At a higher level, a badge can also increase the overall level of user participation on the site. While it may initially seem that these are two different phenomena—participation level versus choice of actions on the site—we can unify them within the model by representing the user's off-site activities via a single additional action type, corresponding to one extra dimension in the space of possible actions. We refer to this additional action type as the *life-action*, since it corresponds to the sum total of the user's activities in his life off the site; the user's preferred distribution of activities will include a probability associated with the life-action as well. With this extended framework, we can ask about the effect of a badge in two distinct respects: the extent to which it draws probability mass away from the life-action, thus increasing participation on the site, and the extent to which is shifts probability mass between different site actions, thus steering behavior within the site.

Empirical Evaluation. After developing the model, we investigate whether its qualitative predictions match the aggregate behavior we see in a large-scale setting that makes strong use of badges. To do this, we analyze the behavior of several million users on Stack Overflow, a large and very active question-answering site. On Stack Overflow, badges play a prominent role—they are displayed alongside a user's name wherever he posts on the site, and they are discussed extensively on Stack Overflow forums.

We find that the main qualitative predictions of the model match the behavior we observe on Stack Overflow in four key respects.

- (i) If a certain level of activity of a particular type is rewarded by a badge, users will increase their activity of this type as they approach the level needed for the badge. This is the notion of *steering* toward a badge boundary mentioned above.
- (ii) Different badges produce different amounts of steering.
- (iii) The extent of steering depends on how close the user is to the badge boundary—an "acceleration" effect in which a user steers particularly strongly toward a badge when they are close to achieving it.
- (iv) Steering involves both shifts in the mixture of actions the user performs on the site, as well as in their overall level of participation on the site.

What we find promising about these results is that each of (i)-(iv) arises naturally as a prediction of our modeling framework, despite the fact that we did not "build in" any of them. Rather, the model generates these behaviors by simply positing users who act to maximize utility.

The Badge Placement Problem. Finally, having proposed a model of user behavior in the presence of badges and showing that its main predictions are consistent with empirical data, we return to the model and explore it from the site designer's perspective. Suppose the site designer has a desired mixture of actions $\mathbf{q}$—over the full user population, she would like to see the action types performed

[1] Throughout this paper we use male pronouns to refer to users of a site and female pronouns to refer to site designers.

in the proportions represented by the distribution $\mathbf{q}$. If the site designer has the ability to define a fixed number of badges, essentially "placing" each one as a frontier in the space of actions, how should she place the badges so that when users optimize their behavior, the overall mixture of activities is as close to $\mathbf{q}$ as possible?

We formalize and study this *badge placement problem*, analyzing the way in which the site designer's solution varies with the properties of the underlying user population and with the desired outcome. From this analysis we can abstract a number of high-level design principles. First, the effectiveness of badges depends significantly on where they are placed, and the optimal location is often surprisingly hard to reach. It is more important to preserve a badge as an incentive over the course of the user's career on the site, even if it is far away for much of that time, than have it be achieved too early. Second, if the site designer wants to use multiple badges to reward the same type of action (*e.g.* one for a medium level of contribution and another for a very high level), the model indicates that it is more useful to spread them out and assign them roughly equal value, rather than concentrating them close together and effectively produce a single badge of greater value. At a qualitative level, these findings suggest the power of creating multiple smaller rewards, relatively far off to preserve their incentive effects over long periods of contributions.

Overall, these results are part of an effort to develop a principled, tractable way of reasoning about badges and their effects on user behavior, validated on evidence from large social sites, and providing a framework for using badges proactively to design incentives that guide a user community toward a desired mix of behaviors.

2. A MODEL OF USER BEHAVIOR

We begin with our theoretical model of user behavior in the presence of badges, following the principles outlined in Section 1. Overall, our goal is to keep the model set-up as simple as possible while still capturing the basic effect of badges as incentives.

2.1 Setting

To establish terminology, a *designer* controls a *site* in which *users* take *actions* of different *types*. Although we talk about the model in terms of user behavior on a Web site, in keeping with our primary motivation, the model is formulated in terms general enough to provide connections back to the off-line examples from Section 1 as well. Moreover, even though we will describe the model from a single user perspective note that our model naturally allows for heterogeneous user populations.

Action types. If there are n possible kinds of actions on the site, we declare the life-action, representing the user's activities outside the site, to be an $(n+1)^{\text{st}}$ action type. The set of possible action types is then denoted $A = \{A_1, A_2, \ldots, A_n, A_{n+1}\}$, where A_{n+1} is the life-action. For example, on a simple question-answering site of the type discussed in the introduction, we might have $n = 4$, with A_1, A_2, A_3, and A_4 corresponding, respectively, to asking a question, answering a question, voting on a question, and voting on an answer; and here A_5 is the life-action.

User histories. A user's history on the site will be represented abstractly as a sequence of choices of action types (*e.g.*, on a Q&A site, the user first asked two questions, then answered a question, then voted on three answers, and so forth). In general, we will be interested in a user's cumulative number of actions taken so far, and these can be represented by a single *action vector* $\mathbf{a} \in \mathbb{R}^{n+1}$ whose i^{th} coordinate $\mathbf{a}^i$ is simply equal to the number of actions of type i the user has taken over his history on the site (with coordinate $\mathbf{a}^{n+1}$ denoting the off-site life-action). We use $\mathbf{e}_i$ to denote the unit vector with a 1 in coordinate i and 0 elsewhere, so that when a user with vector $\mathbf{a}$ performs an action of type i, the user's new vector is $\mathbf{a} + \mathbf{e}_i$.

Badges and their boundaries. We focus on badges that are awarded once the user has reached a certain level of cumulative contribution—so a badge b is associated with the subset of possible user action vectors $\mathbf{a}$ corresponding to contributions that warrant the badge. This set of vectors is monotonic in the sense that if an action vector $\mathbf{a}$ warrants the badge, and $\tilde{\mathbf{a}}$ is coordinate-wise at least as large as $\mathbf{a}$ ($\tilde{\mathbf{a}}^i \geq \mathbf{a}^i$ for all i), then $\tilde{\mathbf{a}}$ also warrants the badge. Note that badges can only be awarded for on-site activities but not for the activity of the life-action. Thus, if $\mathbf{b}$ and $\tilde{\mathbf{b}}$ differ only in coordinate $n+1$ (the life-action coordinate), then $\mathbf{b}$ warrants the badge if and only if $\tilde{\mathbf{b}}$ does.

We write $I_b(\mathbf{a})$ for the indicator function specifying whether vector $\mathbf{a}$ warrants badge b: $I_b(\mathbf{a}) = 1$ if $\mathbf{a}$ warrants the badge, and it is 0 otherwise. Often we will talk about a user reaching the *badge boundary* for a badge b; the badge boundary is simply the set of vectors at which the badge may be first awarded. Concretely, the boundary is the set of vectors $\mathbf{a}$ such that $I_b(\mathbf{a}) = 1$ but for which there exists a unit vector $\mathbf{e}_i$ such that $I_b(\mathbf{a} - \mathbf{e}_i) = 0$.

For example, a common type of badge simply rewards a certain number of actions of a single fixed type (*e.g.*, answering at least a certain total number of questions on a Q&A site); this would mean that for some choice of action type i and a particular threshold k, the badge is awarded to any user whose action vector $\mathbf{a}$ satisfies $\mathbf{a}^i \geq k$. The badge boundary here would be the hyperplane $\mathbf{a}^i = k$.

Utilities and incentives. Finally, we describe a model of the utility a user derives from the site, so as to incorporate the idea that badges produce incentives and thus steer behavior. A user's utility has two components: one from taking types of actions he naturally prefers, and one from receiving badges. The tension between these two components is what drives the behavior of the model.

To represent the user's inherent utility from selecting actions, we want a model based on the idea that the user has a preferred mixture of activities (*i.e.*, a preferred proportion in which to do things on and off the site). We therefore say that a user has a fixed "ideal" distribution $\mathbf{p}$ over action types—in the absence of any other incentives he samples from $\mathbf{p}$ in each step to determine his next action. Basing selection on sampling rather than a sequence of fixed deterministic choices leads to a model whose global structure is very similar to what we would get from deterministic choice, but whose local structure nicely "smooths out" some of the artifacts that would arise at a local level from deterministic selection (for example strict round-robin alternation and parity effects).

In the presence of badges, a user may choose a different mixture of activities, but still has an interest in remaining close to his preferred distribution $\mathbf{p}$. We thus give the user the ability to choose any distribution $\mathbf{x}$ from which to sample the next action, and model the user's goal of staying close to his preferred mixture by saying that he incurs a cost of $g(\mathbf{x}, \mathbf{p})$ for choosing $\mathbf{x}$, where $g(\mathbf{x}, \mathbf{p}) = 0$ if and only if $\mathbf{x} = \mathbf{p}$, and $g(\cdot)$ is monotonic in the sense that $g(\tilde{\mathbf{x}}, \mathbf{p}) \geq g(\mathbf{x}, \mathbf{p})$ if $\tilde{\mathbf{x}} - \mathbf{p}$ is coordinate-wise at least as large as $\mathbf{x} - \mathbf{p}$. Thus the higher the deviation of a user from his preferred actions, the higher the cost. (In our analysis, we will use $g(\mathbf{x}, \mathbf{p}) = \|\mathbf{x} - \mathbf{p}\|_2^2$ for concreteness.)

As we will see, a user may want to choose different distributions in different states, based on his action vector so far. We will use $\mathbf{x_a}$ to denote the distribution chosen by the user when he is at action vector $\mathbf{a}$. If there were no badges, the user would choose $\mathbf{x_a} = \mathbf{p}$ for every $\mathbf{a}$, since g is uniquely minimized for this choice of $\mathbf{x_a}$.

Notice that in this scenario the preferred vector $\mathbf{p}$ doesn't depend on $\mathbf{a}$, and hence the user would act the same regardless of the sequence of actions he has taken so far. However our model is very general, and can be easily generalized to incorporate richer user utilities that may depend extensively on the user's state. Two examples include a user who: (a) wants to maintain the proportions in the vector $\mathbf{p}$ globally over his entire lifetime (that is, he wants to choose a distribution at a given vector $\mathbf{a}$ that tries to bring him back toward a point where the proportion of action types performed so far is balanced according to the coordinates of $\mathbf{p}$), and (b) values badges differently depending on which other badges he has already received—this could be incorporated into the value function V.

Next we consider the utility associated with badges. We use B to denote the set of all badges on the site. A badge $b \in B$ confers a value V_b to a user in each step after he receives it. (Alternately, the value of b could be conferred entirely in a "lump-sum payment" in the step when the user receives b; this would lead to a model that is formally equivalent, but expressing the value as being conferred in each step after the badge is received is an easier formalism to work with.) In keeping with the fact that badges produce genuine incentives on the site under consideration, we treat the notion of a badge's value as a primitive in the model. The broader question of what a badge's value means is an interesting issue that we return to in the discussion at the end of the paper.

In addition to the incentive to achieve a badge, we also want to capture the idea that it is better to receive a badge sooner rather than later. We do this by the standard approach of discounting the future: we say that a user has a fixed exogenous probability $\delta > 0$ that he permanently leaves the system after each action, and write $\theta = 1 - \delta$ for the corresponding probability that he survives to perform another step. The user ceases to receive utility once he leaves the system. In this way, under a plan that involves receiving a badge some distance into the future, the value of the badge is discounted by the probability that he survives long enough in the system to actually reach the badge.[2] We emphasize that the parameter θ controls an exogenous process by which the user leaves the system, in the sense that it operates independently of the user's choices. There is a separate endogenous way for the user to stop performing actions on the site, which is simply to choose a distribution $\mathbf{x}$ over actions that places a probability mass of 1 on the life-action.

Intuitively, at a current vector $\mathbf{a}$, a user therefore needs to choose a distribution $\mathbf{x}$ so as to trade off between remaining close to $\mathbf{p}$ and achieving badges more quickly. This is the trade-off that leads a user to steer toward a badge boundary.

A user's policy. Let us suppose that a user has decided, for each possible action vector $\mathbf{a}$, which distribution $\mathbf{x_a}$ over next actions he will use should he find himself at vector $\mathbf{a}$. We call this choice of distributions $\mathcal{X} = \{\mathbf{x_a}\}$ the user's *policy*. We can write the utility the user in state $\mathbf{a}$ receives from policy $\mathbf{x_a}$, as follows.

$$U(\mathbf{x_a}) = \sum_{b \in B} I_b(\mathbf{a}) V_b + \theta \sum_{i=1}^{n+1} \mathbf{x}_\mathbf{a}^i \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_i}) - g(\mathbf{x_a}, \mathbf{p}) \quad (1)$$

The first sum on the right-hand side is the user's value from all the badges he holds at vector $\mathbf{a}$. The second sum is the expected value arising from the vector the user reaches after sampling according to $\mathbf{x_a}$: with probability $\mathbf{x}_\mathbf{a}^i$ the user moves to $\mathbf{a} + \mathbf{e}_i$, where he will use distribution $\mathbf{x}_{\mathbf{a}+\mathbf{e}_i}$ and hence obtain utility $U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_i})$. This second sum is discounted by θ, since the user only reaches this next step with probability θ. The third term is the cost from the difference between the actual action distribution $\mathbf{x_a}$ and the preferred $\mathbf{p}$.

[2]We focus on this basic form of discounting for simplicity; without much modification one can also formulate the model using more sophisticated types of discounting of the kind developed in the behavioral economics literature [11], and the results in this case are qualitatively unchanged.

2.2 The User's Optimization Problem

Having fully specified our setting, we can now ask how a user will behave under the model in the presence of a set of badges B. Formally, we will assume that the user chooses a policy $\mathcal{X} = \{\mathbf{x_a}\}$ to maximize his utility $U(\mathbf{x_0})$ starting from the origin (*i.e.*, before performing any actions). As with any model of this form, the underlying optimization problem is intended to represent an abstraction of a real user's decision-making, when he trades off the terms in Equation (1), balancing fidelity to his preferred distribution $\mathbf{p}$ against the long-range prospect of obtaining badges.

We note that the user's optimization problem can be cast as the optimum of a Markov decision process (MDP), but this observation doesn't directly help us since MDPs can be computationally expensive to solve even for instances with state sets and action sets of moderate size, and our model here produces instances with a countably infinite state set and a continuum of possible actions. In fact we will be able to develop an efficient algorithm, but it requires making use of the inherent structure of the problem as it arises from our model, rather than invoking a general class of results.

For most of the section, we will focus on *threshold* badges, which reward a certain number of actions of a particular type. Concretely, a threshold badge can be described by a pair (k, i), specifying that a user is awarded the badge as soon as he has taken at least k actions of type A_i (that is, $I_b(\mathbf{a}) = 1$ if and only if $\mathbf{a}^i \geq k$). We say that a threshold badge described by (k, i) *targets* dimension i.

In developing the algorithm, we focus on threshold badges that target one or two dimensions; we will see that such badges illustrate the range of principal behaviors that we subsequently observe in the data as well. At the end of the section, we discuss the extension to general badges.

One targeted dimension

From Equation (1), we see that the user's utility at an action vector $\mathbf{a}$ depends on his utility at each of the vectors $\mathbf{a} + \mathbf{e}_i$. We therefore approach the problem using dynamic programming, with the vectors $\mathbf{a}$ as the *states* of the dynamic program. For simplicity, we will consider the case $n = 2$, so there is one additional site action plus the life-action, but the algorithm trivially extends to larger n.

If we want to set up the dynamic programming recurrence based on Equation (1), then we need to initialize it. We must do this carefully, since each state depends on a state associated with a vector of larger norm. We handle this via the following two observations.

First, note that when the user takes his k^{th} action of type A_1 he achieves the badge, and after this there is no further badge utility to be gained. Hence, for any state $\mathbf{a}$ that is past the badge boundary, the utility from $\mathbf{x_a}$ depends entirely on the cost $g(\mathbf{x_a}, \mathbf{p})$, and so in the optimal policy we have $\mathbf{x_a} = \mathbf{p}$ for all such states. Second, if $\mathbf{a}$ and $\tilde{\mathbf{a}}$ have the same coordinate in dimension 1, then a sequence of actions starting at $\mathbf{a}$ crosses the badge boundary if and only if the same sequence starting at $\tilde{\mathbf{a}}$ crosses the badge boundary. Thus, we effectively have a one-dimensional problem, in which the value of coordinate 1 in the current vector is the only variable that matters.

The utility function $U(\cdot)$ can therefore be expressed simply as a function of $\mathbf{a}^1$, the number of A_1 actions the user has taken. Abusing notation slightly to write U in terms of both vectors and the coordinate in this one targeted dimension, we have

$$\begin{aligned} U(\mathbf{a}^1) &= \theta \sum_{j=1}^{3} \mathbf{x}_\mathbf{a}^j \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_j}) - g(\mathbf{x_a}, \mathbf{p}) \\ &= \theta \cdot [\mathbf{x}_\mathbf{a}^1 U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_1}) + \mathbf{x}_\mathbf{a}^2 U(\mathbf{x_a}) + \mathbf{x}_\mathbf{a}^3 U(\mathbf{x_a})] - g(\mathbf{x_a}, \mathbf{p}) \end{aligned}$$

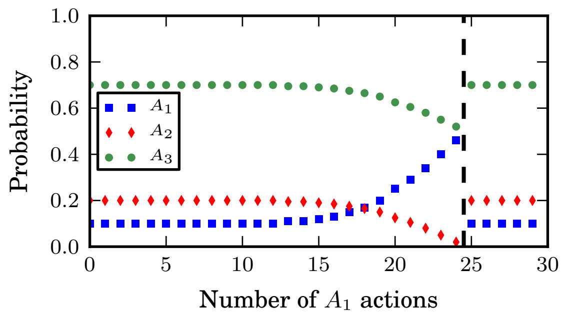

Figure 1: User receives the badge after completing 25 A_1 actions. Actions A_1 increase towards the badge boundary at the expense of the other site action A_2 (shifting effort within the site) as well as the life-action A_3 (increase in site activity).

and then solving for $U(\mathbf{a}^1) = U(\mathbf{x_a})$ we have

$$U(\mathbf{a}^1) = \frac{\theta \cdot \mathbf{x}_\mathbf{a}^1 \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_1}) - g(\mathbf{x_a}, \mathbf{p})}{1 - \theta(\mathbf{x}_\mathbf{a}^2 + \mathbf{x}_\mathbf{a}^3)}$$

Since we have already computed $U(\mathbf{a}^1 + 1) = U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_1})$, this becomes an optimization problem in 3 variables:

$$\begin{aligned} \underset{\mathbf{x_a}}{\text{maximize}} \quad & \frac{\theta \cdot \mathbf{x}_\mathbf{a}^1 \cdot C - g(\mathbf{x_a}, \mathbf{p})}{1 - \theta(\mathbf{x}_\mathbf{a}^2 + \mathbf{x}_\mathbf{a}^3)} \\ \text{subject to} \quad & \mathbf{x}_\mathbf{a}^j \geq 0, \ j = 1, 2, 3 \ \text{ and } \ \sum_{j=1}^{3} \mathbf{x}_\mathbf{a}^j = 1 \end{aligned}$$

where we've replaced $U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_1})$ with C. In the appendix of the extended version of the paper we show how to solve this problem efficiently. For our purposes, the important point is that the optimal distribution in state $\mathbf{a}^1$ can be computed using the solution of the state $\mathbf{a}^1 + 1$. Since we know $\mathbf{x_a} = \mathbf{p}$ for all states $\mathbf{a}$ such that $\mathbf{a}^1 \geq k$, we can use this to compute the optimal $\mathbf{x_a}$ for all $\mathbf{a}$ such that $\mathbf{a}^1 = k - 1$, and recurse all the way back to $\mathbf{a_0}$, thus solving the user's optimization problem in the one-dimensional case.

To illustrate the effects captured by our model, we compute the optimal policy on a simple illustrative instance. In this instance, we place one threshold badge on action A_1 with boundary 25 (*i.e.*, $b = (25, 1)$). We then solve the user optimization problem defined above and plot the optimal $\mathbf{x_a}$ as a function of $\mathbf{a}^1$ (see Figure 1). The user's optimal mixing probability gets progressively more deflected away from his preferred $\mathbf{p}$ as he approaches the badge boundary. The user also increases his probability on A_1 by offloading probability mass from both other action types: he is shifting his effort within the site (moving probability mass from the other site action A_2) and also increasing participation on the site overall (moving probability mass from the life-action A_3).

Note that unifying participation and shifting site effort in this way does not necessarily make them *equivalent* in our framework. The deviation penalty function $g(\mathbf{x_a}, \mathbf{p})$ could be chosen to penalize deviating from user's preference for the life-action more than deviating from his preferences for the various site actions.

Multiple badges. So far we considered a case where there is only one badge on the site. We now show that a similar algorithm can solve the user's optimization problem when there are multiple badges that all target the same dimension.

Let $B = \{b_j = (k_j, 1)\}$ for $j \in \{1, \ldots, m\}$ be a set of m badges all on the same action A_1, and assume that $k_1 < k_2 < \ldots < k_m$ without loss of generality. (If $k_j = k_{j'}$ for some $j \neq j'$, then we can consider these two badges as a single badge with value equal to the sum of their individual values.)

The fact that there are many badge boundaries does not affect the algorithm; the observation about the utility in all states with the same number of A_1 actions still holds, since all such states are equidistant from all badge boundaries, and thus the optimization problem is the same in all of them. The problem thus becomes one-dimensional and solvable in the exact same way as before. The value of each state is "initialized" with $\sum_{b \in B} I_b(\mathbf{a}) V_b$ and again our dynamic programming base case is that in all states $\mathbf{a}$ after all the badge boundaries, the user will choose $\mathbf{x_a} = \mathbf{p}$. Then the region between the last and second-last badge boundaries is identical to the one-badge case we solved in the previous section and can be solved analogously. In general, the region between badges $j - 1$ and j is identical to the single badge case with a badge of value $V_{b_j} + U(\mathbf{x}_{k_j})$ In this way, we recurse backwards through the set of badges to solve the one-dimensional case with many badges.

Two targeted dimensions

Now we consider the case where different badges target different types of actions. We start with $B = \{b_1 = (k_1, 1), b_2 = (k_2, 2)\}$, so there are two dimensions with one badge targeting each (again, let $n = 2$ for convenience).

We begin by observing that only actions on targeted dimensions affect the optimization problem in any state, thus the utility values in two states with the same number of A_1 actions and A_2 actions are the same. Our problem, and corresponding dynamic programming table, is thus two-dimensional. The badge boundaries $\mathbf{a}^1 = k_1$ and $\mathbf{a}^2 = k_2$ split the action space into four regions:

- R: a finite rectangle bounded by the origin and $(k_1 - 1, k_2 - 1)$,
- H: an infinite horizontal strip with boundary points $(k_1, 0)$ and $(k_1, k_2 - 1)$ extending rightward,
- V: an infinite vertical strip with boundary points $(0, k_2)$ and $(k_1 - 1, k_2)$ extending upward, and
- Q: a quadrant rooted at (k_1, k_2).

Similarly to before, past all the badge boundaries the user has no incentive to deviate from $\mathbf{p}$, so $\mathbf{x_a} = \mathbf{p}$ for all states in quadrant Q (those with $\mathbf{a}^1 \geq k_1$ and $\mathbf{a}^2 \geq k_2$).

Quadrants H and V are then identical to the case of one threshold badge in one targeted dimension that we solved above.

Now we are left with the finite rectangle R, which we can directly fill in in order of decreasing coordinate sum since the cells furthest from the origin depend on the value of states we already know from solving quadrants Q, H and V. For every state $\mathbf{a} \in R$:

$$U(\mathbf{x_a}) = \theta \sum_{j=1}^{n+1} \mathbf{x}_\mathbf{a}^j \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_j}) - g(\mathbf{x_a}, \mathbf{p})$$

Consider a state $\mathbf{a}$ in region R that we process in order. We have already computed $U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_1})$ and $U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_2})$, so we can further simplify:

$$U(\mathbf{x_a}) = \frac{\theta \cdot (C_1 \cdot \mathbf{x}_\mathbf{a}^1 + C_2 \cdot \mathbf{x}_\mathbf{a}^2) - g(\mathbf{x_a}, \mathbf{p})}{1 - \theta \cdot \mathbf{x}_\mathbf{a}^3}$$

where $C_j = U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_j})$ for $j = 1, 2$ are the constants we have computed. This results in an optimization problem very similar to the one we had in the one-dimensional case:

$$\begin{aligned} \underset{\mathbf{x_a}}{\text{maximize}} \quad & \frac{\theta \cdot (C_1 \cdot \mathbf{x}_\mathbf{a}^1 + C_2 \cdot \mathbf{x}_\mathbf{a}^2) - g(\mathbf{x_a}, \mathbf{p})}{1 - \theta \cdot \mathbf{x}_\mathbf{a}^3} \\ \text{subject to} \quad & \mathbf{x}_\mathbf{a}^j \geq 0, \ j = 1, 2, 3 \ \text{ and } \ \sum_{j=1}^{3} \mathbf{x}_\mathbf{a}^j = 1 \end{aligned}$$

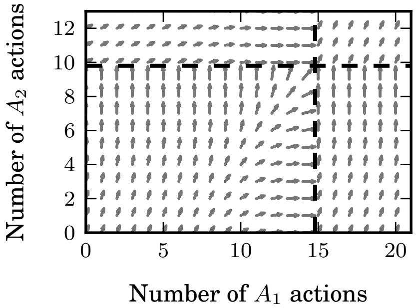

Figure 2: Two badges, each in one dimension. Arrows are projections of the optimal 3-dimensional directions in each state. Arrow length corresponds to user participation on the site, and arrow direction corresponds to the distribution over the two site actions. Notice that towards the badge boundary arrows get longer (as the user increases site activity) as well as point increasingly towards the boundary (as the user shifts effort onto the activity warranting the badge).

Similarly to before, this optimization problem can be solved efficiently. Whereas a naive, approximate solution to the problem requires solving a quadratic program for each possible discretized choice of $\mathbf{x}^3_\mathbf{a}$, our method efficiently finds the exact solution via a reduction to a one-variable optimization problem (details in the appendix of the extended version). This means we can use dynamic programming to fill in R moving right-to-left and top-to-bottom, which solves the user optimization problem with two badges on different targeted dimensions. Figure 2 shows the optimal policy for an example two-dimensional problem. Again, the user both participates more on the site and steers toward badge boundaries.

This algorithm extends to the case when both targeted dimensions have multiple badges, just as in the one-dimensional case. The base case Q region remains as before, the infinite strips H and V become equivalent to the one-dimensional case with many badges, and the finite rectangle R is the same as before except we initialize each state with a different $\sum_b I_b(\mathbf{a})V_b$, depending on which badges have been won. This fully covers the two-dimensional case with any set of threshold badges.

General Monotone Badges. We've seen that interesting and intuitively natural behavior emerges even when we limit our attention to threshold badges in one or two dimensions. We now briefly discuss the user's optimization problem in the case of an arbitrary monotone badge. (We leave the proofs to the extended version.)

For a badge b, let S_b be the set of action vectors that warrant the badge: $S_b = \{\mathbf{a} : I_b(\mathbf{a}) = 1\}$. If we let $m = n + 1$ be the number of dimensions, then S_b is a subset of $\mathbb{N}^m$, the set of all m-tuples of natural numbers, and we say that S_b is *monotone* if for any two vectors $\mathbf{a}, \tilde{\mathbf{a}} \in \mathbb{N}^m$ such that $\tilde{\mathbf{a}} \geq \mathbf{a}$ and $\mathbf{a} \in S$, we have $\tilde{\mathbf{a}} \in S$. In this notation, b is a monotone badge if S_b is monotone.

The first challenge in discussing an arbitrary monotone badge b is how to even specify it compactly—do the conditions required to obtain b necessarily have a finite description? In fact they do, based on a result in combinatorics known as *Dickson's Lemma*. To formulate this result, we need one more definition: for a monotone set $S \subseteq \mathbb{N}^m$, we say that $\mathbf{a} \in S$ is a *minimal element* if for all $\tilde{\mathbf{a}} \neq \mathbf{a}$, we have that $\tilde{\mathbf{a}} \leq \mathbf{a}$ implies $\tilde{\mathbf{a}} \notin S_b$. Dickson's Lemma says the following: any monotone subset of $\mathbb{N}^m$ has only finitely many minimal elements.

This result implies that for any monotone badge b, there is a finite set of vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$ such that $I_b(\mathbf{a}) = 1$ if and only if one of the vectors $\mathbf{c}_i$ satisfies $\mathbf{c}_i \leq \mathbf{a}$.

Given this, we can solve the user's optimization problem for badge b as follows. We first choose a natural number w such that all coordinates of all vectors $\mathbf{c}_i$ are upper-bounded by w, and let $\mathbf{w}$ be the vector in $\mathbb{N}^m$ all of whose coordinates are w. The central construction is to divide up $\mathbb{N}^m$ by thinking of $\mathbf{w}$ as the "origin," and looking at the 2^m regions defined by whether each coordinate is larger or smaller than w. Specifically, let $[m]$ denote the set of dimensions $\{1, 2, \ldots, m\}$; for a set $\sigma \subseteq [m]$, let $T_\sigma \subseteq \mathbb{N}^m$ denote the set of vectors $\mathbf{a} = (\mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^m)$ such that $\mathbf{a}^i \leq w$ if and only if $i \in \sigma$. Notice that $T_{[m]}$ is a finite set, and all other T_σ for $\sigma \neq [m]$ are infinite.

We solve for the optimal policy $\{\mathbf{x_a}\}$ in each region T_σ inductively, in order of increasing cardinality of σ. When σ is the empty set, we know that all $\mathbf{a} \in T_\sigma$ satisfy $\mathbf{x_a} = \mathbf{p}$. For general T_σ, we can prove the following key fact: Let $\mathbf{a}, \tilde{\mathbf{a}} \in T_\sigma$ agree on all coordinates in σ, and let $\mathbf{b} \in \mathbb{N}^m$; then $\mathbf{a} + \mathbf{b} \in S_b$ if and only if $\tilde{\mathbf{a}} + \mathbf{b} \in S_b$. In other words, any action sequence starting from $\mathbf{a}$ leads to obtaining the badge if and only if the same action sequence starting from $\tilde{\mathbf{a}}$ does. For such $\mathbf{a}$ and $\tilde{\mathbf{a}}$, it follows that we can set $\mathbf{x_a} = \mathbf{x}_{\tilde{\mathbf{a}}}$ in the optimal policy.

As a result, it is enough to compute the optimal policy on the subset $T^*_\sigma \subseteq T_\sigma$ in which all coordinates are bounded by $w + 1$; every other $\mathbf{a} \in T_\sigma$ agrees with some $\tilde{\mathbf{a}} \in T^*_\sigma$ on all coordinates in σ, and so its optimal policy can be determined from the optimal policies on T^*_σ. Since T^*_σ is a finite set, we can then compute $\mathbf{x_a}$ for each $\mathbf{a} \in T^*_\sigma$ using the standard recurrence for the optimal policy by considering vectors $\mathbf{a}$ in order of decreasing coordinate sum.

3. EMPIRICAL EVALUATION

Having developed a model of user behavior in the presence of badges, we now investigate how the predictions of the model compare with the aggregate behavior of people on a popular web site. We analyze the question-answering (Q&A) site Stack Overflow, which makes extensive use of badges and was one of the first sites to use them on a large scale.

There are over 100 different badges on Stack Overflow, which vary greatly in how difficult they are to achieve. For instance, there are a number of badges for encouraging new users that nearly everyone obtains, such as the "Editor" badge for contributing a first edit. In order to measure how user behavior changes in relation to the badge, we limit our attention to badges that both require substantial effort to achieve and for which a user's progress towards the badge can be readily determined. This latter restriction excludes one-shot badges such as the "Great Question" badge, which rewards a single high-quality contribution.

Our particular focus is on threshold badges (those that are awarded once a user has taken a prespecified number of actions of certain types) because the action count provides a direct measure of progress towards the badge. There are four main action types on Stack Overflow: asking a question, answering a question, voting on a question and voting on an answer (abbreviated as Q, A, Q-vote and A-vote, respectively). The two badges that we consider here are the "Electorate" badge, awarded for taking at least 600 Q-votes and having at least one Q-vote for every four A-votes[3], and the "Civic Duty" badge, awarded after voting 300 times (on questions or answers). Importantly, users can see exactly how many actions they've taken so far, and how many more they need to take in order to achieve these badges. In the language of our model, Electorate is a threshold badge and Civic Duty is an additive badge.

[3]In our data once user takes 600 Q-votes, they almost always satisfy the restriction on the ratio of Q- to A-votes, so we can think of the "Electorate" badge as being a threshold on 600 Q-votes.

Our analysis is made possible by the granular level of detail available in the activity traces on Stack Overflow[4]. Each individual action performed by a user is recorded and timestamped, which affords us the ability to directly observe the complete sequence of actions users take and measure their progress towards obtaining badges. We use Stack Overflow data from the site's inception on July 31, 2008 to December 31, 2010.

Activity around the badge boundary. We first examine how users' propensities to take different types of actions vary as they approach the badge boundary. We aim to analyze both how users shift their effort between actions on the site and change their overall level of site activity. For each user we bin the number of actions of each type by day. This way, changes in the relative number of various types of actions per day indicate a user shifting his efforts on the site, and the daily sum over all actions measures his overall participation level (where an increase or decrease in participation can be interpreted as steering away from or towards the life-action).

For each badge, we take the complete set of users who ever achieved that badge and axis-align their activity profiles by letting "day 0" denote the day they receive the badge. To eliminate possible population effects, we restrict the set of users to those who were active at least 60 days before and after they win the badge. Figure 3 shows user activity in the days surrounding the awarding of the Electorate and Civic Duty badges. Notice how activity on the targeted actions (Q-votes for Electorate, Q-votes and A-votes for Civic Duty) increases substantially before users achieve the badge, and then almost immediately returns to near-baseline levels. Also notice that most of the other site actions are not adversely affected—the rates of these actions remain relatively stable over time. Since the four actions shown in the Figure are the main activities on the site, this means users increase their overall activity level on the site in the days leading up to achieving these badges. The one exception is the A-vote curve in the Electorate badge, which drops in the days leading up to the badge boundary. This is evidence that users are steering their behavior from A-votes to Q-votes during this time.

Turning towards the badge. Given that we do indeed see the sort of steering behavior predicted by our model, where user activity increases near a badge boundary (in these cases at the expense of the life-action), we now examine *how* users steer towards badge boundaries. One of the main qualitative predictions of our model is that users will "turn" towards badge boundaries, meaning they will deviate more from their preferred actions as they get closer to receiving a badge. We test this prediction by computing which actions a user has taken over the course of his lifetime, and examining how actions change as a function of position in the action space.

We proceed as follows. For every state in the action space, we compute the empirical distribution over site actions that users took in that state. For example, for all users who at one point during their lifetimes had contributed exactly 11 questions, 17 answers, 20 question-votes, and 11 answer-votes, we calculate the distribution over the next action they are going to take. The resulting distribution represents the aggregate direction users traveled at that point in the action space. The composition of these directions forms a vector field like the one we modeled in Figure 2. Since this vector field is 4-dimensional, we visualize its projection onto the question-vote and answer-vote dimensions in Figure 4 (top), where hotter colors represent higher likelihoods of the next action being a question-vote (and cooler colors represent higher likelihoods of the next action being an answer-vote).

[4]Stack Overflow generously gave us a complete trace of actions, but qualitatively similar results are derivable from the data that is publicly available on the Stack Overflow website.

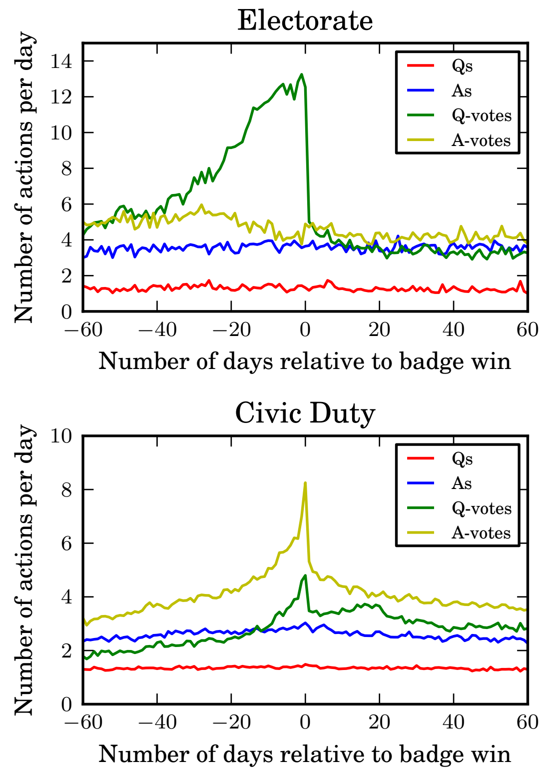

Figure 3: Number of actions per day as a function of number of days relative the time of obtaining a badge. Notice steering in the sense of increased activity on actions targeted by the badge.

The first salient feature of the vector field is the gradient from hot to cold as the angle departing the origin varies between the two extremes. The fact that the color stays the same along a given direction starting from the origin is a validation of our modeling assumption that users have preferred distributions over the action types: one interpretation consistent with this gradient is that users tend to travel in the direction they have already traveled in.

To more clearly illustrate the "turning" effect, we normalize out the tendency of users to act as they have acted in the past by subtracting off the direction of each cell (so that each cell shows the difference between the empirical fraction of actions users chose and the fraction given by the position of the cell). In the resulting plot, Figure 4 (bottom), white indicates a direction matching the vector corresponding to the position in action space, black indicates a higher probability of taking a question-vote, and red indicates a lower probability of taking a question-vote.

The dominant white color for small x values shows that away from badge boundaries, users do not deviate much from the direction they have already taken. For greater x values, we clearly see that once users get near the Electorate badge boundary, they start performing more question votes (than we would expect given their position in the action space). Furthermore, this shifting towards question votes intensifies as they approach the boundary (the darkest greys occur right before the boundary). As soon as users obtain the badge, they shift back away from question voting (doing this action less than we would expect given their position, as indicated by the red color). That the badge boundary naturally emerges from observing users' directions in action space is a striking confirmation that badges influence user behavior and supports our multi-dimensional modeling framework. Additionally, the increase in intensity as we move along the x axis agrees with our model's qualitative prediction that users are increasingly incentivized as they approach badge boundaries.

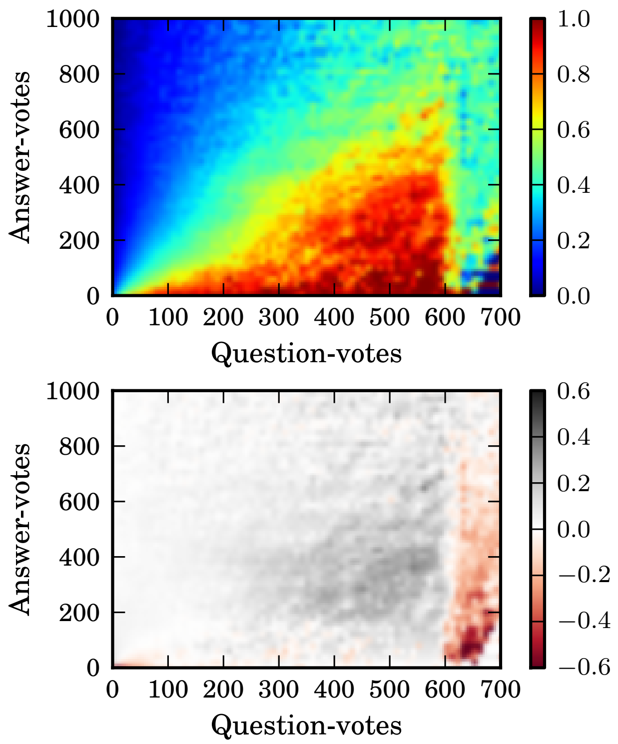

Figure 4: Electorate badge. Given that user has taken x question-votes and y answer-votes, what is the probability that next action will be a question vote. Top: Raw probability. Bottom: Relative change in probability of question-voting. Notice the effects of "turning" towards the badge boundary.

To summarize, the above experiments demonstrate that badges can steer user behavior in a manner consistent with our model. First, we observe user site activity increases as users approach the badge and thus site activity increases at the expense of the life-action. And second, we observe that users also shift their effort on the site towards actions that lead them to badges. To our knowledge, this is the first concrete evidence of users changing their behavior in either of these two distinct ways in response to badges.

4. THE BADGE PLACEMENT PROBLEM

Having seen that our model matches up well with real-world behavior for threshold badges on Stack Overflow (badges that are awarded once a user has taken a specified number of actions of certain types) we now investigate how our model can help provide insights into the design of badges for online communities. In particular, the site designer has the ability to decide on the conditions for obtaining badges; within our model we think of this as "placing" the badge boundaries within the space of action vectors. We are interested in addressing the following questions: (1) How much user steering do different badge placements provide? (2) How might site designer place badges to best achieve desired user behavior? (3) What is the space of user behaviors fixed badges can elicit?

We present our results organized around these three central questions. First, we find that the effectiveness of a badge is generally maximized at an internal optimum that is surprisingly high, and we explore how the effectiveness varies with the setting. Second, the effectiveness of multiple badges working together is maximized when they are of equal value and are spaced roughly evenly apart. Finally, we explore the "feasible region" of user behaviors the designer can elicit with a fixed set of badges and find it to be a highly complex and counterintuitive structure worthy of future study.

We arrived at these results by running experiments on our model developed in Section 2. By repeating the simulations using many different parameter settings, we found that the presented results all hold for a wide range of parameter values and thus illustrate broader principles of the badge placement problem. Moreover, we used the same $g(\cdot)$ as before: $g(\mathbf{x}, \mathbf{p}) = \|\mathbf{x} - \mathbf{p}\|_2^2$.

Throughout this section, we refer to the total fraction of actions on a targeted action (over the user's lifetime) that results from a particular set of badge placements as the *yield* of those badges. We also define *gain* to be the difference between the yield and the default fraction of actions the user takes in the absence of any badges.

Optimal location and yield with one badge

We start our investigations with the case in which the site designer has a single badge at her disposal and wishes to maximize the yield on a single dimension, which we take to be A_1 without loss of generality. This setting, while simple, is of both conceptual and practical value, since the effects of placing a single badge are most easily seen when the badge is the only influence over user behavior, and in practice it is often desirable for the designer to limit herself to a single badge.

Placing a badge of fixed prespecified value on a single dimension requires striking the right balance between two mutually competing forces. First, if the badge is to have an effect on many actions, its threshold should be set high enough so that it takes many steps to achieve it. However, if its threshold is set too high, then even surviving long enough on the site to achieve the badge becomes a low-probability event and thus users will not be sufficiently incentivized to steer strongly towards it. The solution to the badge placement problem is therefore in general an internal optimum between these two goals.

Figure 5 shows how the fraction of actions on the targeted dimension A_1 (the yield) varies as a function of where the badge of fixed value is placed. We plot this for various choices of $\mathbf{p} = (x, 0.5 - x, 0.5)$, $x \in \{0.05, \ldots, 0.45\}$ (while keeping all other parameters fixed). Notice that each curve has a unique internal maximum. We can also see that the resulting fraction of A_1 actions is higher for larger values of $\mathbf{p}^1$, which intuitively makes sense since the user is increasingly predisposed to take A_1 actions. More subtle, however, is that the optimal badge location also increases with $\mathbf{p}^1$: for $\mathbf{p}^1 = 0.05$ it is $A_1 = 75$, whereas for $\mathbf{p}^1 = 0.45$ it is $A_1 = 90$. Another interesting observation is that this optimal location is surprisingly high: since $\theta = 0.99$ in this example, $\mathbf{p}^1 = 0.05$ implies that the user would only take 5 A_1 actions in the absence of badges, yet the optimal badge location is far beyond this point, at $A_1 = 75$.

These results suggest interesting relationships between the model parameter settings and the amount of steering induced by the badge. For example, we just saw that the yield increases with $\mathbf{p}^1$, and that the badge gets placed further away from the origin as $\mathbf{p}^1$ increases. In Figure 6, we plot the yield at the optimal badge location as a function of $\mathbf{p}^1$, V_b, and the user's expected lifetime $1/(1-\theta)$. We observe that the relationship follows a different functional form for each setting: the yield increases approximately linearly in $\mathbf{p}^1$ (but the gain decreases linearly), increases with diminishing returns in V_b, and surprisingly *decreases* in the expected lifetime of the user. The decreasing relationships are particularly interesting; they imply that the designer can steer users more on the actions they dislike (i.e., low $\mathbf{p}^1$) than those they like, and that badges have stronger effects when users' lifetimes are shorter.

Two principles for placing multiple badges

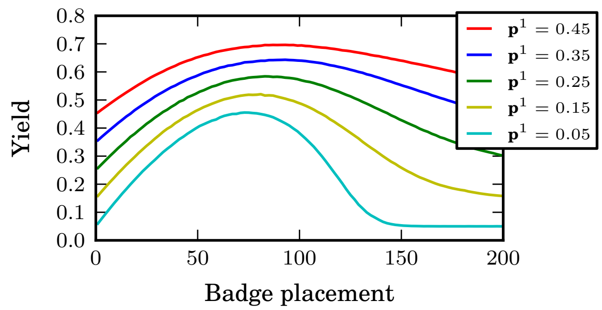

Figure 5: The resulting fraction of actions on the targeted dimension (here A_1) as a function of where the badge is placed. The different curves show how the relationship varies as the user's preferences p change.

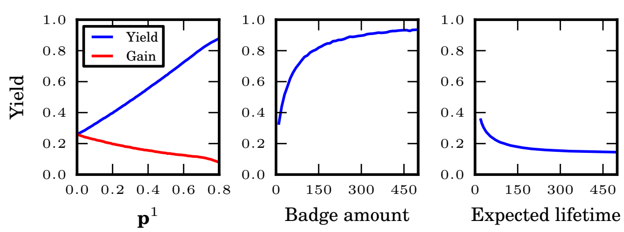

Figure 6: How optimal yield varies with setting parameters.

So far we have established that a single badge is optimally placed at an ideal middle ground (which is surprisingly high) and explored how the optimal yield depends on the setting. Now we investigate how one should place use multiple badges in concert with each other (*i.e.*, target the same dimension). Exploring optimal placements for pairs of badges (rather than individual badges) uncovers two additional basic principles that we now discuss.

Optimal spacing. We find that yield on a single targeted dimension is maximized by two badges when they are spaced roughly evenly apart. In Figure 7, each point (x, y) represents placing the two badges at x and y A_1 actions, respectively, and the color indicates the resulting yield (i.e., the cumulative activity on both targeted actions). More effective badge placements are colored red, while less effective placements are colored blue. Due to symmetry the plot is naturally mirrored below the diagonal.

The salient qualitative observation is that the optimal set of badge locations is off the main diagonal, indicating that it is better to place the two badges at distinct locations than to combine them into a single large badge at any single location.

In fact, this observation is consistent with the solution of the user's optimization problem in Section 2.2 in the case of multiple badges on the same dimension, where we saw that the last badge exerts influence over not only the region after the second-last boundary, but also over the earlier regions as well. There, once we solved the user's optimal direction for the last badge, solving the second-last badge region became equivalent to the one-badge case with a badge value equal to the second-last badge value *plus* the expected utility from potentially winning the last badge.

We further note that in the optimal set of placements, badges are spaced approximately evenly apart. If the first badge is at $\mathbf{a}^1 = x$, the second badge is at $\mathbf{a}^1 = 2x - \epsilon$ for some small ϵ. The reason for the ϵ is that the second badge is placed a little closer to the origin than it would be placed in a one-badge problem starting at $\mathbf{a}^1 = x$ since it not only steers behavior for $\mathbf{a}^1 \geq x$ but also has a slight effect on the user for $\mathbf{a}^1 < x$.

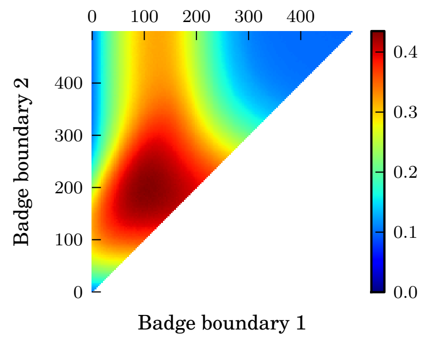

Figure 7: Yield as a function of where two badges are placed on the same targeted dimension. Every cell (x, y) corresponds to placing one badge at $\mathbf{a}^1 = x$ actions and another at $\mathbf{a}^1 = y$ actions, and the color of the cell indicates the resulting yield.

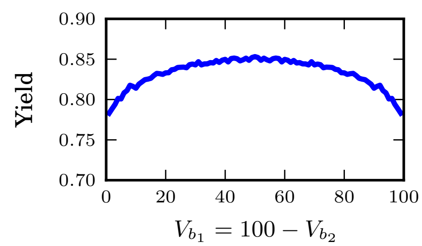

Figure 8: How the optimal yield achievable with two badges depends on how a fixed amount of utility is split across the two badges. The more even the split, the higher the yield.

Yield is maximized with badges of equal value. In the previous case, the two badges targeted the same action but both had some fixed equal value z. Now we allow the two badges to take different values while their total value is held constant. Assume the first badge has value x and the second $2z - x$. Now fix z and the question is: what is the optimal way to split the value of $2z$ among the two badges to gain maximum yield? Figure 8 plots the yield as a function of x. Notice the effectiveness of the two badges is maximized when the badges have equal value ($x = z$). This result suggests that the designer should create badges in such a way that they have about equal value.

Targeting two dimensions

We just established how yield on one targeted dimension is maximized by a single badge and when two badges are acting together. The yield of a set of badges was a natural objective function, since it expresses how much of the given action can be induced by the badge(s). With badges that can target different dimensions, there is a more complex set of trade-offs, since we need to decide how we want to balance increases in the volume of one site action with increases in another, when both are targeted by badges. More generally, we ask: what is the set of all possible behaviors that can be induced using badges targeting multiple dimensions?

For simplicity we consider the scenario in which the designer has two threshold badges that she can place however she wants. She can put them both on A_1, both on A_2, or one on each action. Figure 9 plots the all possible behaviors that can be induced using two threshold badges. The general contour of the plot is maintained across a wide range of model parameter settings. (For example, the contour grows as the user's expected lifetime gets longer.)

The opacity of each (x, y) cell reflects how many distinct placements of two badges produce a yield of x on A_1 and a yield of y on A_2. The dark outlines of the blue and red regions (where the designer targets the same dimension with both badges) indicate that it's easier (in the sense that there are more unique placements) to

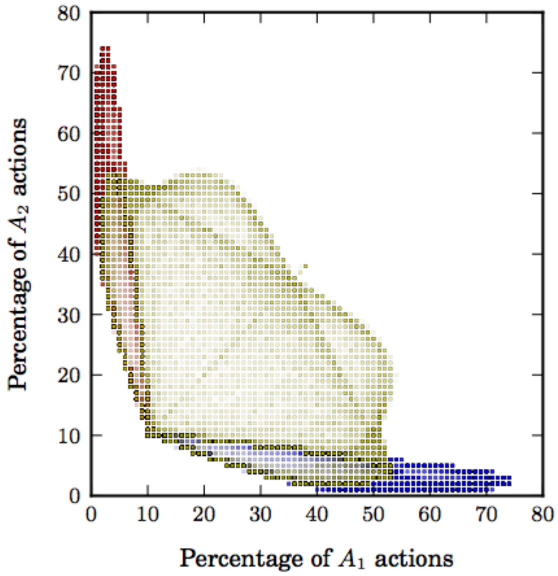

Figure 9: The feasible region of what yields are achievable with two badges of fixed value (here $V_{b_1} = V_{b_2} = 200$ and $\mathbf{p} = (0.1, 0.1, 0.8)$). A blue cell at (x, y) indicates that the designer can have both badges target A_1 actions and elicit $x\%$ of A_1 actions and $y\%$ of A_2 actions; red cells are similar but for the A_2 dimension; and yellow cells indicate what is achievable by putting one badge on each dimension. The cell opacity corresponds to how many different pairs of badge locations elicit that particular yield.

elicit yield on the border than in the middle of these wing-shaped regions. If we consider the blue region as an example, the upper outline, near the $y = 10$ line, is dark because it corresponds to using one badge nearly optimally and essentially "discarding" the other one (by placing it so far from the origin that it doesn't shape user behavior), whereas the lower dark outline (smoothly varying from $(10\%, 10\%)$ to $(40\%, 2\%)$) corresponds to using both badges in concert. There are similar recurring "veins" in the gold region, indicating that some parameterized lines through the feasible region are implemented by more badge placements than others.

The maximum yield on a single dimension naturally occurs when the designer places both badges on that dimension, and the middle ground where both the yield is high on both dimensions occurs when she places one badge on each dimension. However, note that the feasible region is interestingly non-convex between these two cases. This means that designer cannot induce certain behaviors regardless of how she places the badges. For example, the designer can elicit $(5\%, 70\%)$ yield using two badges on A_2 and $(15\%, 50\%)$ using one badge on each, but she cannot elicit the intermediate $(10\%, 60\%)$ using any combination of two badges.

5. RELATED WORK

As noted in the introduction, the use of badges is a growing trend in the design of online communities, social computing applications, and electronic commerce (*e.g.* see [19]), and researchers have begun to study the role that badges play in these sites. Antin and Churchill [1] present a conceptual organization for different types of badges, considering, among other things, the motivation they provide from a social-psychological viewpoint. Oktay et al. [18] use quasi-experimental designs to support causal claims in the context of badges. The use of badges can be viewed as part of the growing phenomenon of *gamification* [9], in which elements traditionally associated with computer games are used to motivate people in other domains. Our work contributes to this literature by proposing a formal framework for reasoning about the effects badges will have, which can potentially be used for the badge design in both current and new contexts.

Badges and badge-like recognitions have been proposed in offline domains as well. There is a growing movement toward using systems of badges in educational settings; as one example, the Badges for Lifelong Learning Competition has made use of the Mozilla Open Badges project [10, 16] for this purpose. Thus far such initiatives in education have used badges primarily as a form of credentialing, but moving beyond this towards engaging and motivating students is another key opportunity with badges in this setting. The advantages of badges in such contexts is a theme that has been advocated in earlier research in education as well [4].

Rewards for cumulative effort have been considered in several domains. For example, one can interpret customer loyalty programs as containing badge-like incentives, such as the different status milestones in airline frequent flier programs, and research economics and marketing has studied the effect of such programs on customer behavior [14, 15]. In a different direction, Zhang et al. consider placing limited amounts of reward in a Markov decision process, as an instance of what they term *environment design* [20].

Finally, our work is related to the more general question of incentives for contribution in social media. This is a very broad area, encompassing a number of approaches beyond just the use of badges for expressing incentives. As noted in the introduction, two methodologies that have been brought to bear on this question in the computing literature are (i) social-psychological perspectives on the notion of engagement and social motivators [5, 6, 17]; and (ii) algorithmic game-theoretic and economic approaches, including incentives for recruitment, contribution quality, and crowdsourced effort [2, 7, 8, 12, 13].

6. CONCLUSION

Badge systems are an increasingly widespread feature of online social media sites, and they can produce strong incentive effects on the users in these domains. Our work has proposed a methodology for reasoning about these incentives, starting with a model of users who optimize their behavior given opportunities to receive badges. The main qualitative predictions of our model are borne out by an analysis of user behavior in the presence of badges on Stack Overflow, and we have seen how the model provides useful qualitative insights into the problem of placing badges to optimize their incentive effects.

We see a number of important directions in which further work could be pursued. First, the notion of a badge's *value* has been taken as a primitive definition in this work, but as discussed earlier, it is interesting to think about the mechanisms by which badges acquire (non-monetary) value, and the extent to which a site designer can influence the value of a badge. More generally, our model suggests that in optimizing a system of badges, certain parameters—not just the badges' values but also the structure of the possible actions, users' preferences for these actions, and users' expected lifetimes—all can play a potential role in the process. Developing methods of estimating these parameters for use in the design of badge systems is an interesting direction for further research.

Incentivizing users to increase their activity naturally brings up the question of how this affects the quality of their actions. For example, we find on Stack Overflow that users' votes on questions are significantly more positive before they receive the Electorate badge than after it. Developing principled ways of incorporating action quality into models of user behavior in the presence of badges is an exciting direction for future work.

Acknowledgments. We thank Stack Overflow for providing their data and Hristo Spassimirov Paskov for help solving the optimization problems. This research is supported in part by a Google PhD Fellowship, a Simons Investigator Award, a Google Research Grant, an ARO MURI grant, the Okawa Foundation, Boeing, Allyes, Intel, a Sloan Fellowship, a Microsoft Faculty Fellowship, and NSF grants IIS-1149837, IIS-0910664, CCF-0910940, IIS-1159679, IIS-1016909, CNS-1010921, and IIS-1016099.

7. REFERENCES

[1] J. Antin and E. Churchill. Badges in social media: A social psychological perspective. In *CHI Workshop on Gamification*, 2011.

[2] M. Babaioff, S. Dobzinski, S. Oren, and A. Zohar. On bitcoin and red balloons. In *ACM Conference on Electronic Commerce*, pages 56–73, 2012.

[3] T. Becker and V. Weispfenning. *Gröbner Bases: A Computational Approach to Commutative Algebra*. Springer, 1993.

[4] J. H. Bishop. Incentives for learning: Why american high school students compare so poorly to their counterparts overseas. Technical Report CAHRS 89-09, Cornell University School of Industrial and Labor Relations, 1989.

[5] M. Burke, C. Marlow, and T. M. Lento. Feed me: motivating newcomer contribution in social network sites. In *ACM Conference on Human Factors in Computing Systems*, pages 945–954, 2009.

[6] M. Burke and B. Settles. Plugged in to the community: social motivators in online goal-setting groups. In *International Conference on Communities and Technologies*, pages 1–10, 2011.

[7] M. Cebrián, L. Coviello, A. Vattani, and P. Voulgaris. Finding red balloons with split contracts: robustness to individuals' selfishness. In *ACM Symposium on Theory of Computing*, pages 775–788, 2012.

[8] S. Chawla, J. D. Hartline, and B. Sivan. Optimal crowdsourcing contests. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 856–868, 2012.

[9] S. Deterding, M. Sicart, L. Nacke, K. O'Hara, and D. Dixon. Gamification: Using game-design elements in non-gaming contexts. In *CHI Workshop on Gamification*, 2011.

[10] Digital Media and Learning Competition. Badges for lifelong learning. http://dmlcompetition.net/competition/4/badges-competition-cfp.php.

[11] S. Frederick, G. Loewenstein, and T. O'Donoghue. Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401, 2002.

[12] A. Ghosh and R. P. McAfee. Incentivizing high-quality user-generated content. In *International World Wide Web Conference*, pages 137–146, 2011.

[13] J. Kleinberg and P. Raghavan. Query incentive networks. In *IEEE Symposium on Foundations of Computer Science*, pages 132–141, 2005.

[14] P. K. Kopalle and S. A. Neslin. The economic viability of frequency reward programs in a strategic competitive environment. *Review of Marketing Science*, 1, 2003.

[15] M. Lewis. The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research*, 41(3), 2004.

[16] Mozilla. Open badges. http://openbadges.org/en-US/.

[17] H. L. O'Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6):938–955, 2008.

[18] H. Oktay, B. Taylor, and D. Jensen. Causal discovery in social media using quasi-experimental designs. In *Proceedings of the First Workshop on Social Media Analytics*, pages 1–9. ACM, 2010.

[19] N. Singer. You've won a badge (and now we know all about you). *New York Times*, 4 February 2012.

[20] H. Zhang, D. C. Parkes, and Y. Chen. Policy teaching through reward function learning. In *ACM Conference on Electronic Commerce*, pages 295–304. ACM, 2009.

APPENDIX

A. GENERAL MONOTONE BADGES

At the end of Section 2.2, we briefly discussed the user's optimization problem in the case of an arbitrary monotone badge. Here we give an extended argument with proofs showing that our algorithm extends to this general setting.

Let $\mathbf{N}^m$ denote the set of all m-dimensional vectors whose coordinates are natural numbers. We say $S \subseteq \mathbf{N}^m$ is *monotone* if for any two vectors $\mathbf{a}, \tilde{\mathbf{a}} \in \mathbf{N}^m$ such that $\tilde{\mathbf{a}} \geq \mathbf{a}$ and $\mathbf{a} \in S$, we have $\tilde{\mathbf{a}} \in S$. A vector $\mathbf{a}$ is a *minimal element* of S if (i) $\mathbf{a} \in S$, and (ii) if $\tilde{\mathbf{a}} \neq \mathbf{a}$ satisfies $\tilde{\mathbf{a}} \leq \mathbf{a}$ then $\tilde{\mathbf{a}} \notin S_b$.

We begin with a theorem from combinatorics known as *Dickson's Lemma.*

> *(1) (Dickson's Lemma). Let $S \subseteq \mathbf{N}^m$ be monotone. Then S has only finitely many minimal elements.*

It will also be useful to describe a particular method for finding a minimal element of a monotone set $S \subseteq \mathbf{N}^m$ starting from any $\mathbf{a} \in S$. The methods is as follows. Fix any permutation π on the dimensions $\{1, 2, \ldots, m\}$. Starting from $\mathbf{a}$, we decrease the value of $\mathbf{a}$'s coordinate in dimension $\pi(1)$ until the last point at which we still have a vector in S. Call the vector we reach at this point $\mathbf{a}_1 \in S$. We then do the same starting from $\mathbf{a}_1$ using dimension $\pi(2)$, ending at a vector $\mathbf{a}_2 \in S$. We continue through all m dimensions this way according to π, ending with a vector $\mathbf{a}_m \in S$; we denote $\mathbf{a}_m$ by $\pi(\mathbf{a})$.

> *(2) For any monotone $S \subseteq \mathbf{N}^m$, any $\mathbf{a} \in S$, and any permutation π, the point $\pi(\mathbf{a})$ is a minimal element of S.*

To prove (2), suppose by way of contradiction that $\pi(\mathbf{a})$ is not minimal; this means that for some coordinate i, we have $\pi(\mathbf{a}) - \mathbf{e}_i \in S$, where $\mathbf{e}_i$ denotes the unit vector in coordinate i. Now, in the process that led from $\mathbf{a}$ to $\pi(\mathbf{a})$, consider the vector $\tilde{\mathbf{a}}$ at the moment when we had finished working on dimension i. We have $\pi(\mathbf{a}) \leq \tilde{\mathbf{a}}$, and hence $\pi(\mathbf{a}) - \mathbf{e}_i \leq \tilde{\mathbf{a}} - \mathbf{e}_i$. But $\pi(\mathbf{a}) - \mathbf{e}_i \in S$ by assumption, and $\tilde{\mathbf{a}} - \mathbf{e}_i \notin S$ by the definition of the process, and this contradicts the monotonicity of S.

Now we discuss badges. Consider a system where there are m possible actions, and a user has a preferred distribution $\mathbf{p}$ over these actions. The user's current state is a vector $\mathbf{a} \in \mathbf{N}^m$, the set of all m-tuples of natural numbers. (We can think of $m = n + 1$ for a system with n actions plus one life-action.) A badge b has an associated indicator function I_b, where $I_b(\mathbf{a}) = 1$ if the badge has been earned at vector $\mathbf{a}$, and $I_b(\mathbf{a}) = 0$ otherwise. Let $S_b = \{\mathbf{a} : I_b(\mathbf{a}) = 1\}$; we say that the badge b is *monotone* if S_b is a monotone set.

We say that a *point-badge* is a badge b for which there exists $\mathbf{c}$ such that $I_b(\mathbf{a}) = 1$ if and only if $\mathbf{a} \geq \mathbf{c}$. We say that an $\vee$-badge is a badge b that is an "or" of a finite set of point-badges—that is, there exist $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$ such that $I_b(\mathbf{a}) = 1$ if and only if $\mathbf{a} \geq \mathbf{c}_i$ for some $i \in \{1, 2, \ldots, k\}$. In this case, we will say that $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$ is a *generator set* for b.

Now, clearly an $\vee$-badge is a monotone badge. But there is also a non-trivial converse to this observation, as follows.

(3) A badge is monotone if and only if it is an $\vee$-badge.

Since one direction is easy, the crux of proving (3) is to show that any monotone badge is an $\vee$-badge. To prove this, let b be a monotone badge. Since S_b is a monotone subset of $\mathbf{N}^m$, Dickson's Lemma says that it has only finitely many minimal elements; we denote all the minimal elements of S_b by $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$.

Let b' be the $\vee$-badge with generator set $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$. We claim that b and b' are the same badge, in the sense that $S_b = S_{b'}$. Clearly $S_{b'} \subseteq S_b$, since $\mathbf{c}_i \in S_b$ for all i, and S_b is monotone. Conversely, consider any $\mathbf{a} \in S_b$; there must exist a minimal element c_i of S_b for which $\mathbf{c}_i \leq \mathbf{a}$. It follows that $\mathbf{a} \in S_{b'}$, and since $\mathbf{a}$ was arbitrary, we have $S_b \subseteq S_{b'}$. Thus $S_b = S_{b'}$, and this proves (3).

Now, consider a monotone badge b, with a generator set $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$. Suppose that all coordinates of all vectors $\mathbf{c}_i$ are upper-bounded by the natural number w, and let $\mathbf{w}$ be the vector in $\mathbf{N}^m$ all of whose coordinates are w.

The key construction is to divide up $\mathbf{N}^m$ by thinking of $\mathbf{w}$ as the "origin," and looking at the 2^m regions defined by whether each coordinate is larger or smaller than w. Specifically, let $[m]$ denote the set of dimensions $\{1, 2, \ldots, m\}$; for a set $\sigma \subseteq [m]$, let $T_\sigma \subseteq \mathbf{N}^m$ denote the set of vectors $\mathbf{a} = (\mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^m)$ such that $\mathbf{a}^i \leq w$ if and only if $i \in \sigma$. Notice that $T_{[m]}$ is a finite set, and all other T_σ for $\sigma \neq [m]$ are infinite.

The set $T_{[m]}$ has the property that all generators of S_b lie in $T_{[m]}$. This leads to the following crucial property of our decomposition into the sets T_σ.

(4) Let $\mathbf{a}, \tilde{\mathbf{a}} \in T_\sigma$ agree on all coordinates in σ, and let $\mathbf{b} \in \mathbf{N}^m$. Then $\mathbf{a} + \mathbf{b} \in S_b$ if and only if $\tilde{\mathbf{a}} + \mathbf{b} \in S_b$.

Suppose by way of contradiction that there exist $\mathbf{a}$, $\tilde{\mathbf{a}}$, and $\mathbf{b}$ as in the statement of (4) such that $\mathbf{a} + \mathbf{b} \notin S_b$ but $\tilde{\mathbf{a}} + \mathbf{b} \in S_b$. Let $\mathbf{a}_0$ be the coordinatewise maximum of $\mathbf{a}$ and $\tilde{\mathbf{a}}$; by monotonicity, $\mathbf{a}_0 + \mathbf{b} \in S_b$.

Now, we reduce $\mathbf{a}_0 + \mathbf{b}$ to a minimal element of S_b via the procedure described before Statement (2), using a permutation π in which we reduce all coordinates in the complement of σ before any coordinate in σ. Let $\mathbf{a}_*$ denote the vector we reach at the moment when all the coordinates in the complement σ have just been completed. It cannot be the case that $\mathbf{a}_* \leq \mathbf{a} + \mathbf{b}$, since $\mathbf{a}_* \in S_b$ while $\mathbf{a} + \mathbf{b} \notin S_b$. But $\mathbf{a}_*$ and $\mathbf{a} + \mathbf{b}$ agree on all coordinates in σ, and so it follows that $\mathbf{a}_*^i > (\mathbf{a} + \mathbf{b})^i$ for some coordinate $i \notin \sigma$. Finally, $\pi(\mathbf{a}_0 + \mathbf{b})^i = \mathbf{a}_*^i$, since the remainder of the procedure to produce $\pi(\mathbf{a}_0 + \mathbf{b})$ from $\mathbf{a}_*$ never modifies coordinate i again, and hence $\pi(\mathbf{a}_0 + \mathbf{b})^i > (\mathbf{a} + \mathbf{b})^i$. But since $i \notin \sigma$, we have $\mathbf{a}^i > w$ and hence $(\mathbf{a} + \mathbf{b})^i > w$. This contradicts the definition of w—all minimal elements of S_b have all coordinates bounded by w—and hence this completes the proof of (4).

Using (4), we can now describe our dynamic programming approach to the user's optimization problem in the presence of badge b, consisting of a choice of $\mathbf{x_a}$ for each $\mathbf{a} \in \mathbf{N}^m$. We start by setting $\mathbf{x_a} = \mathbf{p}$ for all $\mathbf{a} \in S_b$; since $T_\phi \subseteq S_b$, this includes all $\mathbf{a} \in T_\phi$. We then determine the values of $\mathbf{x_a}$ for $\mathbf{a} \in T_\sigma$ by induction on the cardinality of σ: we can start with $\sigma = \phi$ as the base case, using $\mathbf{x_a} = \mathbf{p}$ for all $\mathbf{a} \in T_\phi$.

In general, consider a set $\sigma \subseteq [m]$, and suppose we have determined the values of all $\mathbf{x_a} \in T_\psi$ for all ψ that are proper subsets of σ. The key point is that if $\mathbf{a}, \tilde{\mathbf{a}} \in T_\sigma$ agree on all coordinates in σ, then (4) implies that $\mathbf{x_a} = \mathbf{x}_{\tilde{\mathbf{a}}}$, since any sequence of actions starting at $\mathbf{a}$ results in the badge if and only if the same sequence of actions starting at $\tilde{\mathbf{a}}$ results in the badge.

Given this, for $\mathbf{a} \in T_\sigma$, let $f(\mathbf{a})$ be the vector in T_σ that agrees with $\mathbf{a}$ on all coordinates in σ, and in which all coordinates not in σ have been set equal to $w + 1$. Let $T_\sigma^* = \{f(\mathbf{a}) : \mathbf{a} \in T_\sigma\}$. We have $\mathbf{x}_{f(\mathbf{a})} = \mathbf{x_a}$, and so it is enough to determine $\mathbf{x_a}$ just for all $\mathbf{a} \in T_\sigma^*$.

The set T_σ^* is finite, since each coordinate of a vector $\mathbf{a} \in T_\sigma^*$ is bounded by $w + 1$. We determine the values of $\mathbf{x_a}$ for all $\mathbf{a} \in T_\sigma^*$ by backward induction on the sum of coordinates in $\mathbf{a}$. When we get to a vector $\mathbf{a}$ in this computation, we can evaluate the utility of choosing a vector $\mathbf{x_a}$ in state $\mathbf{a}$ via the recurrence

$$U(\mathbf{x_a}) = I_b(\mathbf{a})V_b + \theta \sum_{i=1}^{m} \mathbf{x}_\mathbf{a}^i \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_i}) - g(\mathbf{x_a}, \mathbf{p})$$

Of the terms on the right-hand side, the first is a constant and the third has a closed-form expression. In the sum that forms the second term, each summand $\mathbf{x}_\mathbf{a}^i \cdot U(\mathbf{x}_{\mathbf{a}+\mathbf{e}_i})$ involves a vector $\mathbf{a} + \mathbf{e}_i$ such that $f(\mathbf{a} + \mathbf{e}_i)$ satisfies one of three possible properties: (i) it belongs to T_ψ^* for some ψ that is a proper subset of σ; (ii) it belongs to T_σ^* and has larger coordinate sum than $f(\mathbf{a})$; or (iii) $f(\mathbf{a} + \mathbf{e}_i) = f(\mathbf{a})$. In the first two cases we can fill in the value of the summand by induction; in the third case, we have $\mathbf{x_a} = \mathbf{x}_{\mathbf{a}+\mathbf{e}_i}$ in the optimal solution, and so we can bring these summands over the left-hand side. In this way, we can solve a maximization problem in the coordinates of $\mathbf{x_a}$ to find the optimal choice of $\mathbf{x_a}$.

Working backward by induction in this way, we can thus compute the optimal choice of $\mathbf{x_a}$ for all states $\mathbf{a} \in \mathbf{N}^m$.

B. OPTIMIZATION PROBLEMS

In Section 2.2, we claimed that the user's optimization problem in two targeted dimensions can be efficiently solved. Whereas a naive, approximate solution to the problem requires solving many quadratic programs, here we show how to efficiently find the exact solution via a reduction to a one-variable optimization problem. (The technique for solving the optimization problem in one targeted dimension is similar after a suitable change of variables.)

We wish to solve the following optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & \frac{\sum_{j=1}^{n} \theta C^j \mathbf{x}^j - \|\mathbf{x} - \mathbf{p}\|_2^2}{1 - \theta \cdot \mathbf{x}^{n+1}} \\ \text{subject to} \quad & \mathbf{x}^j \geq 0, \ j = 1, 2, \ldots, n+1, \\ & \sum_{j=1}^{n+1} \mathbf{x}^j = 1 \end{aligned}$$

Let $\mathbf{C} = [C^1 \ C^2 \ldots C^n]^T$ be the vector of constants C^j for $j = 1, 2, \ldots, n$, and let $\tilde{\mathbf{x}} = [x^1 \ x^2 \ldots x^n]^T$ and $\tilde{\mathbf{p}} = [p^1 \ p^2 \ldots p^n]^T$ be the projections of $\mathbf{x}$ and $\mathbf{p}$ onto the first n dimensions respectively. If we consider fixing $\mathbf{x}^{n+1} \geq 0$ for the moment, this problem can be recast as the following minimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \|\tilde{\mathbf{p}} - \tilde{\mathbf{x}}\|_2^2 - \theta \cdot \mathbf{C}^T \tilde{\mathbf{x}} \\ \text{subject to} \quad & \tilde{\mathbf{x}}^j \geq 0, \ j = 1, 2, \ldots, n, \\ & \sum_{j=1}^{n} \mathbf{x}^j = 1 - \mathbf{x}^{n+1} \end{aligned}$$

First we add the equality constraint to the objective function $f(\tilde{\mathbf{x}})$ with a Lagrange multiplier λ:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\tilde{\mathbf{x}}) = \|\tilde{\mathbf{p}} - \tilde{\mathbf{x}}\|_2^2 - \theta \cdot \mathbf{C}^T\tilde{\mathbf{x}} + \lambda(\mathbf{1}^T\tilde{\mathbf{x}} - 1 + \mathbf{x}^{n+1}) \\ \text{subject to} \quad & \tilde{\mathbf{x}}^j \geq 0,\ j = 1, 2, \ldots, n \end{aligned}$$

Now, we can expand the norm to express $f(\tilde{\mathbf{x}})$ as

$$f(\tilde{\mathbf{x}}) = \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - 2\tilde{\mathbf{p}}^T\tilde{\mathbf{x}} + \tilde{\mathbf{x}}^T\tilde{\mathbf{x}} - \theta\mathbf{C}^T\tilde{\mathbf{x}} + \lambda(\mathbf{1}^T\tilde{\mathbf{x}} - (1 - \mathbf{x}^{n+1}))$$

By taking partial derivatives and setting them to zero, we can express the solution in terms of λ:

$$\begin{aligned} \frac{\delta f(\tilde{\mathbf{x}})}{\delta x^i} &= -2\tilde{\mathbf{p}}^i + 2\tilde{\mathbf{x}}^i - \theta\mathbf{C}^i + \lambda = 0 \\ &\rightarrow \quad \tilde{\mathbf{x}}^i = \frac{1}{2}[2\tilde{\mathbf{p}}^i + \theta\mathbf{C}^i - \lambda]_+ \end{aligned}$$

where $[x]_+ = \max(x, 0)$ enforces the non-negativity constraint. It will be convenient to define the vector $\mathbf{z}$ such that $\mathbf{z}^i = \mathbf{p}^i + \theta\mathbf{C}^i/2$. For any vector $\mathbf{v}$, let $\mathbf{v}_\lambda$ denote the vector such that $\mathbf{v}_\lambda^i = \begin{cases} \mathbf{v}^i & 2\mathbf{z}^i > \lambda \\ 0 & \text{otherwise} \end{cases}$. Then we can rewrite $\|\tilde{\mathbf{p}} - \tilde{\mathbf{x}}\|_2^2 - \theta\mathbf{C}^T\tilde{\mathbf{x}}$ as:

$$\begin{aligned} &\|\tilde{\mathbf{p}} - \tilde{\mathbf{x}}\|_2^2 - \theta\mathbf{C}^T\tilde{\mathbf{x}} = \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \sum_{i:2\mathbf{z}^i>\lambda}(\tilde{\mathbf{p}}^i(2\mathbf{z}^i - \lambda)) + \frac{1}{4}\sum_{i:2\mathbf{z}^i>\lambda}(2\mathbf{z}^i - \lambda)^2 \\ &\qquad - \frac{\lambda}{2}\sum_{i:2\mathbf{z}^i>\lambda}\mathbf{C}^i(2\mathbf{z}^i - \lambda) \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^T(2\tilde{\mathbf{p}}_\lambda + \theta\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda) \\ &\quad + \frac{1}{4}\|2\tilde{\mathbf{p}}_\lambda + \theta\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda\|_2^2 - \frac{\theta}{2}\mathbf{C}^T(2\tilde{\mathbf{p}}_\lambda + \lambda\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda) \end{aligned}$$

where $\mathbf{1} = [1\,1\,\ldots\,1]^T$ is the vector of all 1s. Now we can expand the norm and use the fact that for any vectors $\mathbf{a}$ and $\mathbf{b}$, $\mathbf{a}^T\mathbf{b}_\lambda = \mathbf{b}^T\mathbf{a}_\lambda = \mathbf{b}_\lambda^T\mathbf{a}_\lambda$ to simplify further:

$$\begin{aligned} &\|\tilde{\mathbf{p}} - \tilde{\mathbf{x}}\|_2^2 - \theta\mathbf{C}^T\tilde{\mathbf{x}} = \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \frac{1}{2}\tilde{\mathbf{p}}^T(2\tilde{\mathbf{p}}_\lambda + \theta\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda) \\ &\quad - \frac{\theta}{4}\mathbf{C}^T(2\tilde{\mathbf{p}}_\lambda + \theta\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda) - \frac{\lambda}{4}\mathbf{1}^T(2\tilde{\mathbf{p}}_\lambda + \theta\mathbf{C}_\lambda - \lambda\mathbf{1}_\lambda) \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^T\tilde{\mathbf{p}}_\lambda - \theta\tilde{\mathbf{p}}^T\mathbf{C}_\lambda - \frac{\theta^2}{4}\mathbf{C}^T\mathbf{C}_\lambda + \frac{\lambda^2}{4}n_\lambda \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \|\tilde{\mathbf{p}}_\lambda + \frac{\theta}{2}\mathbf{C}_\lambda\|_2^2 + \frac{\lambda^2}{4}n_\lambda \\ &= \tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \|\mathbf{z}_\lambda\|_2^2 + \frac{\lambda^2}{4}n_\lambda \end{aligned}$$

where $n_\lambda = |\{i : 2\mathbf{p}^i + \theta\mathbf{C}^i > \lambda\}|$. Finally, we derive $\mathbf{x}^{n+1}$ using the fact that $\mathbf{x}^{n+1} = 1 - 1^T\tilde{\mathbf{x}}$:

$$\begin{aligned} \mathbf{x}^{n+1} &= 1 - \mathbf{1}^T\tilde{\mathbf{x}} \\ &= 1 - \frac{1}{2}\sum_{i:2\mathbf{z}^i>\lambda}(2\tilde{\mathbf{p}}^i + \theta C^i - \lambda) \\ &= 1 - \sum_{i:2\mathbf{z}^i>\lambda}\mathbf{z}^i + \frac{n_\lambda}{2}\lambda \\ &= 1 - \mathbf{1}^T\mathbf{z}_\lambda + \frac{n_\lambda}{2}\lambda \end{aligned}$$

Now we have expressed $\mathbf{x}$ in terms of λ, so we can simply substitute back into our original function:

$$\begin{aligned} &\frac{\theta\mathbf{C}^T\tilde{\mathbf{x}} - \|\tilde{\mathbf{x}} - \tilde{\mathbf{p}}\|_2^2 - (\mathbf{x}^{n+1} - \mathbf{p}^{n+1})^2}{1 - \theta\mathbf{x}^{n+1}} = \\ &= \frac{-(\tilde{\mathbf{p}}^T\tilde{\mathbf{p}} - \|\mathbf{z}_\lambda\|^2 + \lambda^2 n_\lambda/4) - (1 - \mathbf{1}^T\mathbf{z}_\lambda + \lambda n_\lambda/2 - \mathbf{p}^{n+1})^2}{1 - \theta(1 - \mathbf{1}^T\mathbf{z}_\lambda + \lambda n_\lambda/2)} \end{aligned}$$

We have now transformed the problem to an equivalent one-dimensional optimization problem. The final expression is a piecewise quadratic-over-linear function of λ with breakpoints at $2\mathbf{z}^i$ for $i = 1, 2, \ldots, n$. We solve each piece with elementary calculus and take the max over all pieces, which is then the global optimum of our original problem.