

Steering User Behavior with Badges

Fonte: *Steering User Behavior with Badges*, A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec, WWW 2013, Rio De Janeiro, Brazil

per
Sistemi Complessi
Laurea Magistrale in Informatica
Università di Bologna
AA 2013/2014

di
Mario Taddei
mario.taddei2@studio.unibo.it

Cos'è un Badge?

- In generale, un badge è un distintivo che viene visualizzato per indicare un risultato speciale o come simbolo di autorità.
- Nel videogaming, un badge è un achievement.
- Nel nostro caso, un badge è un riconoscimento dato ad un utente per dei contributi ad un sito.

Indice

- Modello formale sul comportamento di un utente in presenza di badge
- Studio dell'uso di badge e del loro effetto sul sito Stack Overflow
- Studio del problema di posizionamento di un badge

“Give me enough medals and I'll win you any war” - Napoleone

Modello del comportamento di un utente

Definiamo:

- n : possibili tipi di azioni
- *life-action*: $(n+1)$ -esimo tipo di azioni
- $A = \{A_1 + A_2 + \dots + A_n + A_{n+1}\}$ = insieme dei tipi di azioni
- a : vettore delle azioni (esempio: $[2, 3, 2, 1, 1]$)
- e_i : versore (esempio: $e_2 = [0, 1, 0, 0, 0]$)
- B : insieme di tutti i badge su un sito
- b : badge
- $I_b(a)$: funzione booleana, indica se il vettore a conferisce il badge b
- N.B.
- Il *badge boundary* è il confine che decreta l'acquisizione del badge. $I_b(a) = 1$ AND $I_b(a - e_i) = 0$

Modello del comportamento di un utente

User Utility: è composta da 2 componenti, una è fare le azioni preferite e un'altra è fare le azioni che portano ad un badge.

La tensione tra le due componenti guida il modello.

Definiamo:

- p : distribuzione di probabilità di scelta delle prossime azioni preferita dell'utente (esempio $[0.4, 0.1, 0.2, 0.2, 0.1]$)
- x : distribuzione di probabilità di scelta delle prossime azioni
- $g(x, p)$: costo per cambiare la distribuzione dalla preferita a x
- x_a : distribuzione quando il vettore è a

n : num tipi
 $n+1$: life-action
 A : insieme tipi
 a : vettore azioni
 e_i : versore
 B : insieme badge
 b : badge
 $I_b(a)$: ho badge?

Modello del comportamento di un utente

Definiamo:

- V_b : valore conferito dal badge
- δ : probabilità che un utente se ne vada per sempre (valutata ad ogni azione)
- $\theta = 1 - \delta$: probabilità che un utente rimanga

N.B:

- Un badge non può essere ottenuto tramite life-action
- L'utente cessa di ricevere utilità quando lascia il sistema
- Il valore del badge tiene conto di tutte le probabilità che fanno rimanere l'utente nel sistema

n: num tipi
n+1: life-action
A: insieme tipi
a: vettore azioni
 e_i : versore
B: insieme badge
b: badge
 $I_b(a)$: ho badge?
p: dist preferita
x: dist generica
g(x,p): func costo
 x_a : dist con vett a

Modello del comportamento di un utente

$$U(x_a) = \sum_{b \in B} I_b(a) V_b + \theta \sum_{i=1}^{n+1} x_{ai} \cdot U(x_a + e_i) - g(x_a, p)$$

- La prima parte indica la somma di tutti i valori dei badge che l'utente ha al vettore a .
- La seconda indica la somma delle utilità di tutte le prossime azioni, considerando la probabilità che l'utente rimanga.
- La terza indica il costo nell'allontanarsi dalla distribuzione preferita.

n : num tipi
 $n+1$: life-action
 A : insieme tipi
 a : vettore azioni
 e_i : versore
 B : insieme badge
 b : badge
 $I_b(a)$: ho badge?
 p : dist preferita
 x : dist generica
 $g(x,p)$: func costo
 x_a : dist con vett a
 V_b : valore badge
 θ : prob user resta

Modello del comportamento di un utente

Definiamo:

- threshold badge = coppia (k,i) , indica che mancano k azioni di tipo i per acquisire il badge

N.B.

- Una volta ottenuto il badge, la distribuzione delle azioni sarà nuovamente p
- k definisce le chiamate ricorsive che servono per calcolare la user utility

n : num tipi
 $n+1$: life-action
 A : insieme tipi
 a : vettore azioni
 e_i : versore
 B : insieme badge
 b : badge
 $I_b(a)$: ho badge?
 p : dist preferita
 x : dist generica
 $g(x,p)$: func costo
 x_a : dist con vett a
 V_b : valore badge
 θ : prob user resta

Modello del comportamento di un utente

Consideriamo un badge su una dimensione, e assumiamo $n=2$ per semplicità.
Threshold = (25,1)

N.B.

- La k -esima azione di tipo 1 garantisce il badge
- Se $a_1 = a'_1$, allora una sequenza di azioni che garantisce il badge partendo da a , garantirà lo stesso badge partendo da a' .

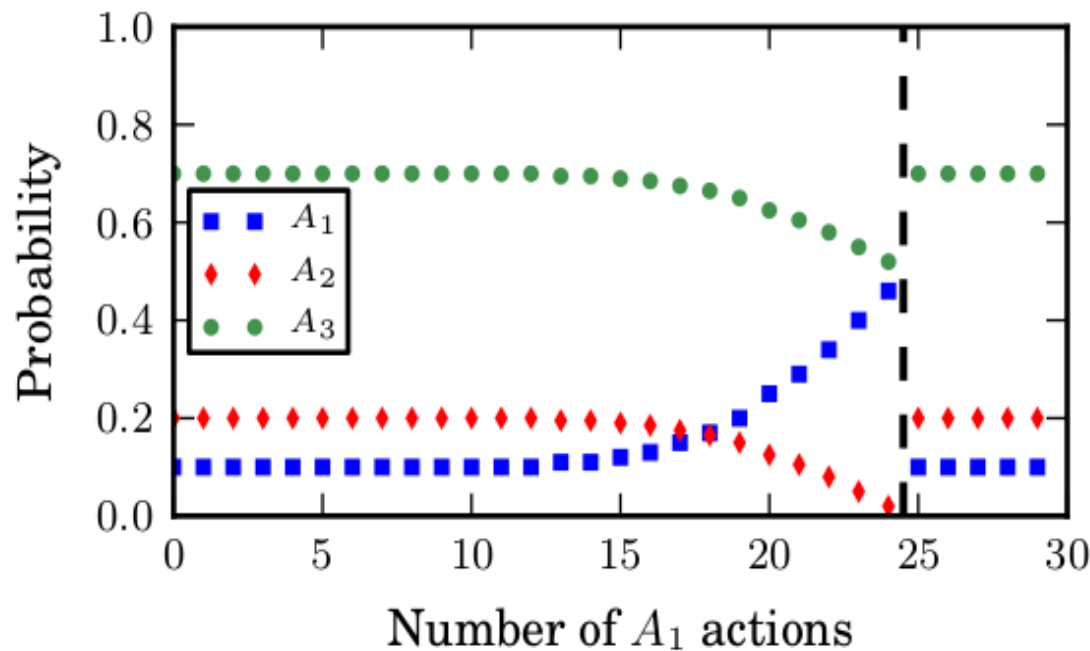
n : num tipi
 $n+1$: life-action
 A : insieme tipi
 a : vettore azioni
 e_i : versore
 B : insieme badge
 b : badge
 $I_b(a)$: ho badge?
 p : dist preferita
 x : dist generica
 $g(x,p)$: func costo
 x_a : dist con vett a
 V_b : valore badge
 θ : prob user resta
 (k,i) =threshold

Modello del comportamento di un utente

$$U(a_1) = \frac{(\theta x_{a1} \cdot U(x_a + e_1) - g(x_a, p))}{(1 - \theta(x_{a2} + x_{a3}))}$$

$$\text{maximize } x_a \frac{(\theta x_{a1} \cdot C - g(x_a, p))}{(1 - \theta(x_{a2} + x_{a3}))}$$

n: num tipi
 n+1: life-action
 A: insieme tipi
 a: vettore azioni
 e_i: versore
 B: insieme badge
 b: badge
 I_b(a): ho badge?
 p: dist preferita
 x: dist generica
 g(x,p): func costo
 x_a: dist con vett a
 V_b: valore badge
 θ: prob user resta
 (k,i)=threshold

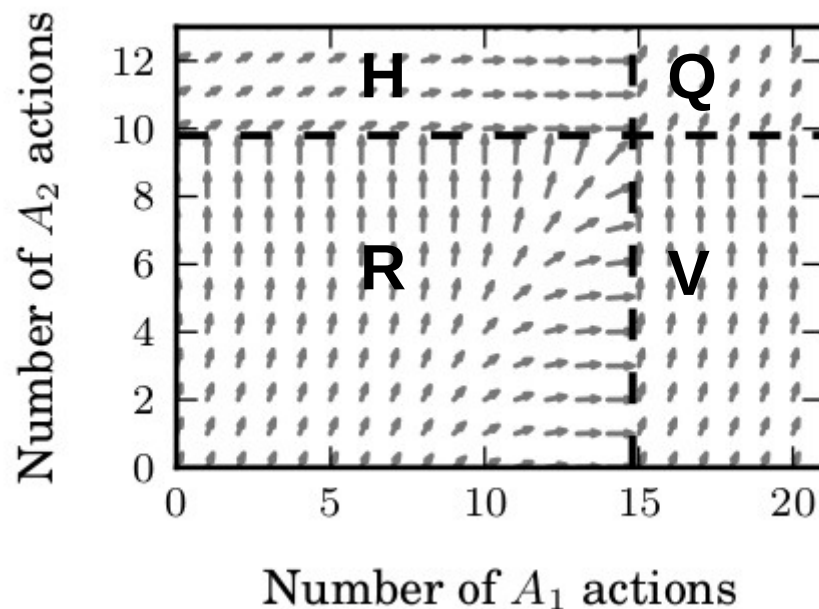


Modello del comportamento di un utente

Consideriamo due badge su due dimensioni, e assumiamo $n=2$.

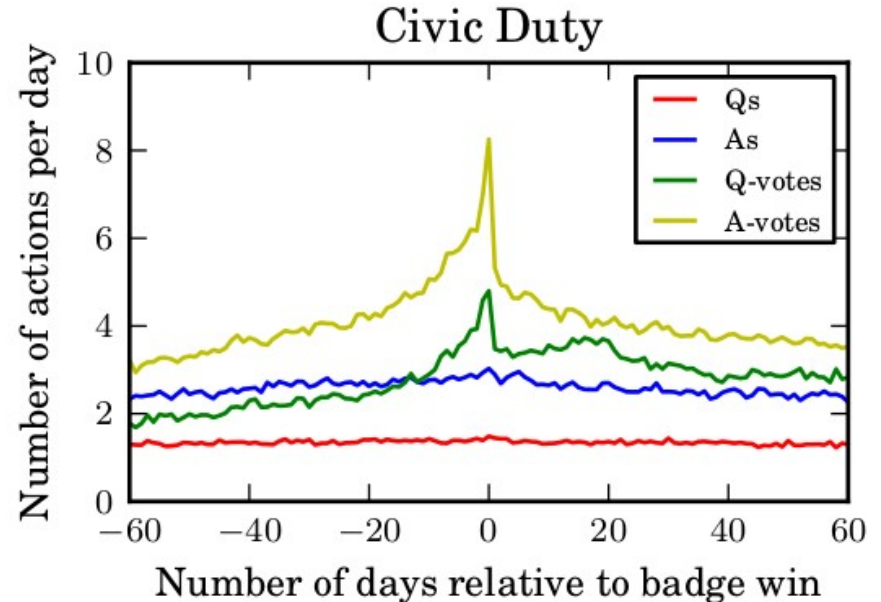
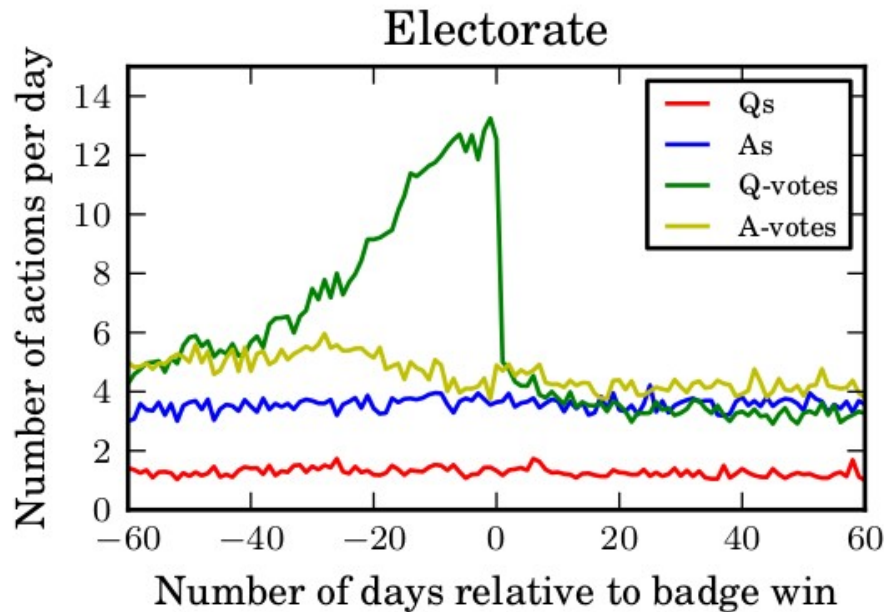
Thresholds= $[(15,1),(10,2)]$

$$\text{maximize } x_a \frac{(\theta(C_1 \cdot x_{a1} + C_2 \cdot x_{a2}) - g(x_a, p))}{(1 - \theta(x_{a3}))}$$



n : num tipi
 $n+1$: life-action
 A : insieme tipi
 a : vettore azioni
 e_i : versore
 B : insieme badge
 b : badge
 $I_b(a)$: ho badge?
 p : dist preferita
 x : dist generica
 $g(x,p)$: func costo
 x_a : dist con vett a
 V_b : valore badge
 θ : prob user resta
 (k,i) =threshold

Badge su StackOverflow



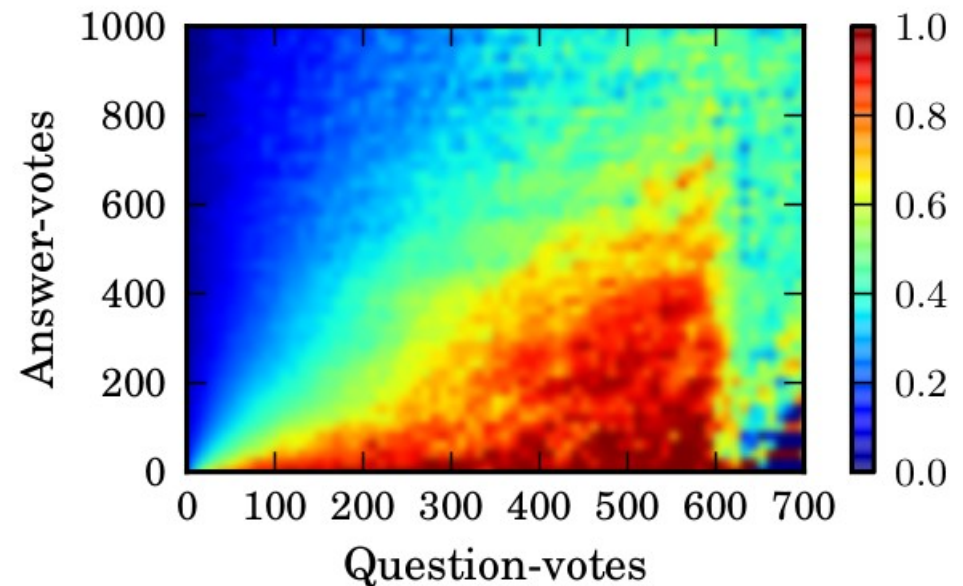
- Electorate badge: fare almeno 600 Q-votes (monodimensionale)
- Civic Duty badge: fare almeno 300 Q-votes o A-votes (bidimensionale)

Badge su StackOverflow

- Per ogni stato nello spazio delle azioni è stata calcolata empiricamente la distribuzione delle azioni fatte dagli utenti.
- Esempio: per ogni utente che si è trovato ad avere 11q, 17a, 20qv e 11av si è calcolata la distribuzione sull'azione successiva
- Il grafico riporta solo le due dimensioni qv e av
- Il rosso indica alta probabilità di effettuare q-vote

N.B.

- Gli utenti hanno una distribuzione preferita sui tipi di azione
- Gli utenti tendono a mantenere la direzione che hanno scelto, in particolare quando sono lontani dal raggiungimento del badge
- Una volta vicini al confine dell'Electorate badge gli utenti tendono a fare molte azioni q-vote



Posizionare un badge

Due forze contrastanti:

- Fare in modo che il badge abbia effetto su più azioni.
- Fare in modo che non siano troppe altrimenti sarebbe poco probabile trovare un utente che faccia tutti i passi per raggiungerlo.

Definiamo:

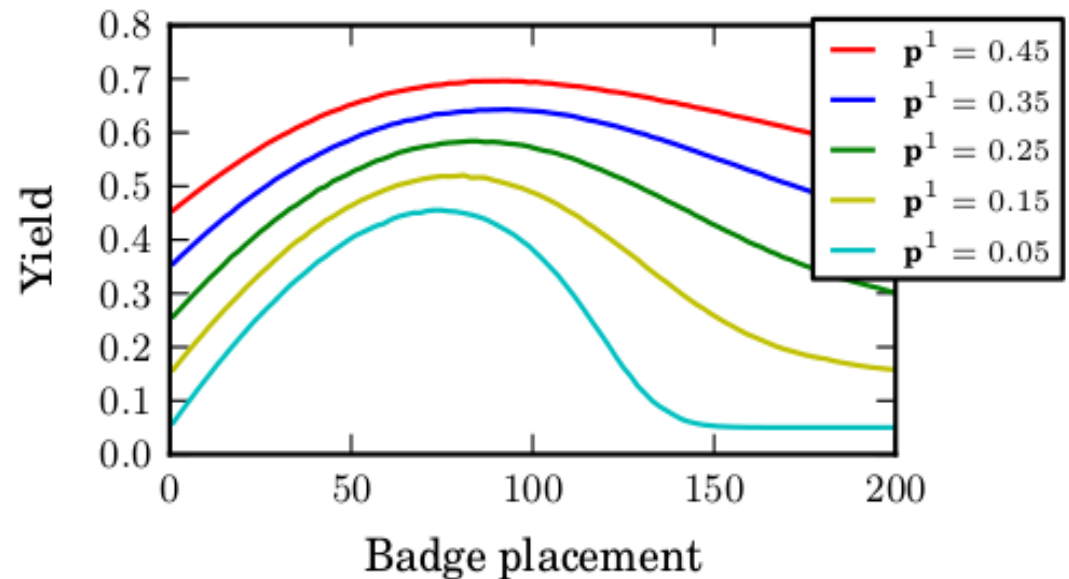
- Yield: il rapporto tra le azioni di una specifica dimensione e le azioni di tutte le dimensioni

Posizionare un badge

Il grafico mostra che secondo lo yield, le posizioni migliori sono quelle più impegnative per un utente

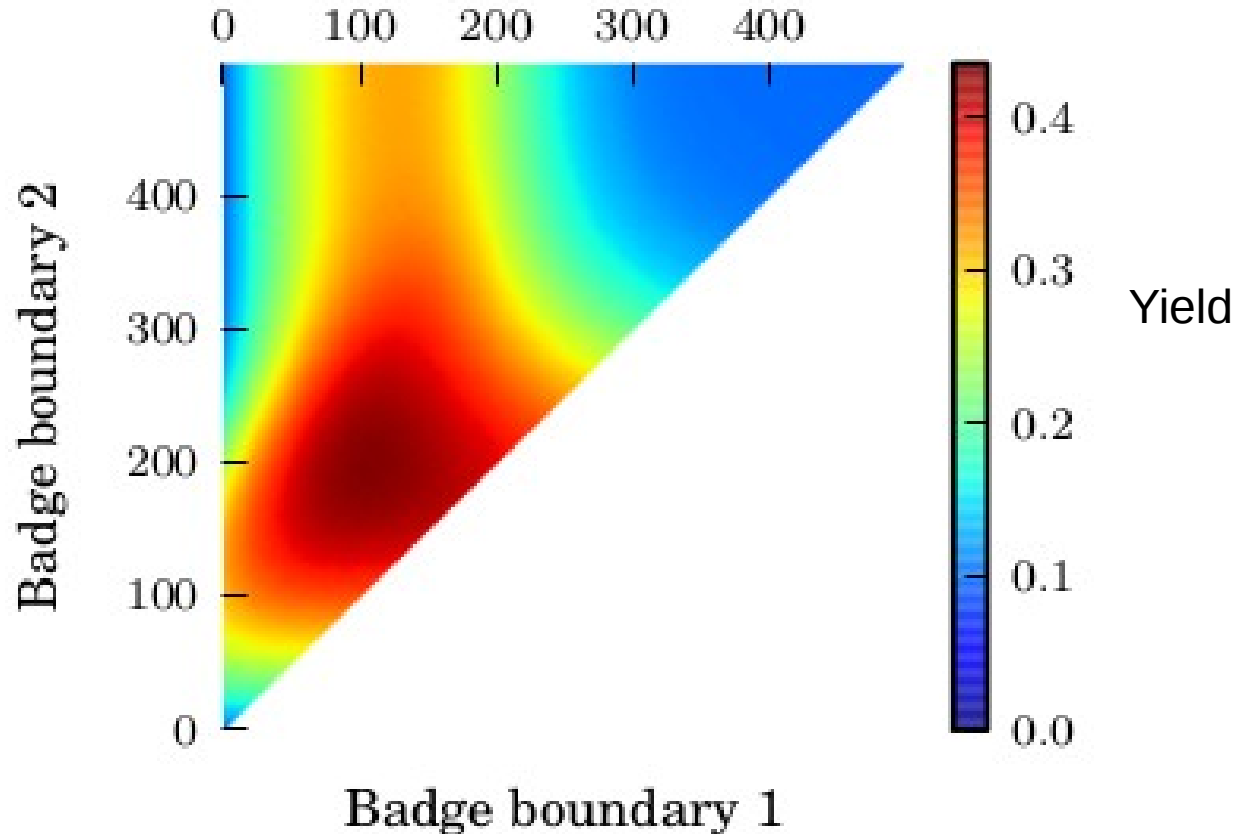
$$p = (x, 0.5 - x, 0.5), x \in [0.05, \dots, 0.45]$$

Con $\theta = 0.99$ l'utente farebbe solamente 5 azioni sulla dimensione considerata, ma la posizione ottimale del badge è 75



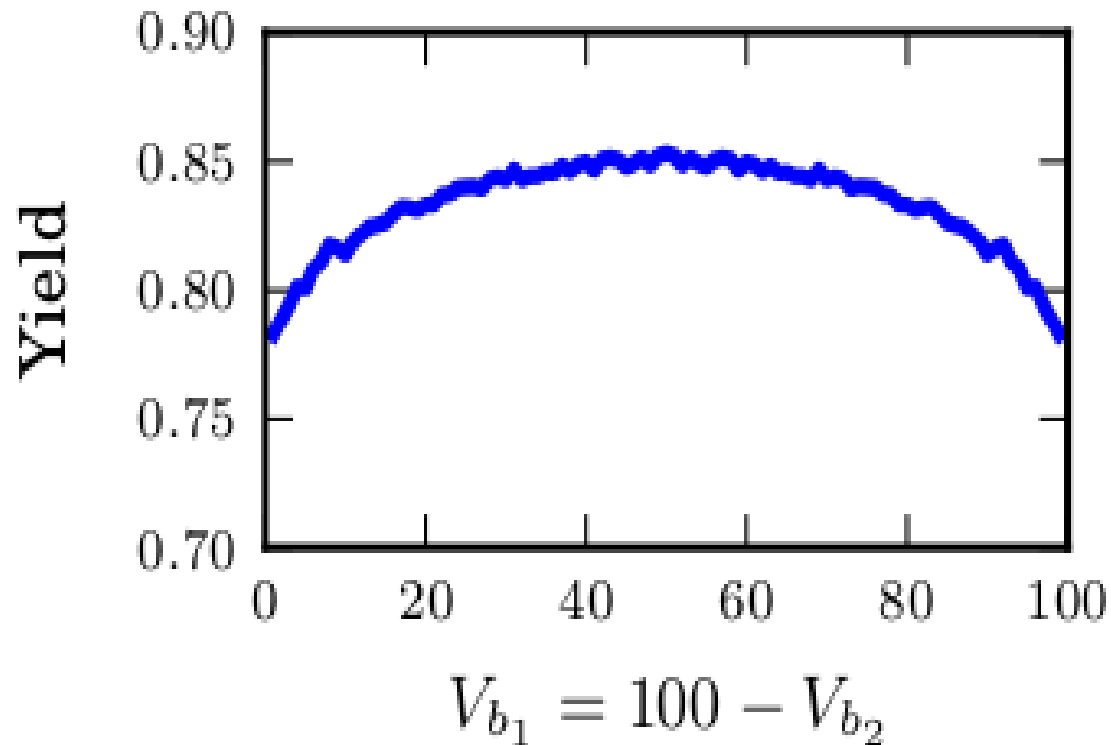
Posizionare un badge

Sulla stessa dimensione, i badge devono essere posizionati alla stessa distanza



Posizionare un badge

Nel caso di due badge sulla stessa dimensione, si ottiene massimo yield dando loro lo stesso valore.



Posizionare un badge

Nel caso di due badge in due dimensioni, i badge possono essere posizionati su dimensioni diverse o sulla stessa.

Nel grafico i punti più opachi indicano le possibili posizioni dei badge in base allo yield.

L'intensità rossa indica lo yield per le azioni di tipo A1, blu per A2.

Un designer può scegliere la combinazione (5,70), o (15,50), ma non (10,60).

