The Hiring Problem and Lake Wobegon Strategies

Ravi Kumar
Yahoo! Research
Sunnyvale, CA
ravikumar@yahoo-inc.com
Joint work with

• Andrei Broder (Yahoo! Research)
• Adam Kirsch (Harvard)
• Michael Mitzenmacher (Harvard)
• Eli Upfal (Brown)
• Sergei Vassilvitskii (Yahoo! Research)
You know the Google story: small start-up of highly-skilled programmers in a garage grows into a large international company. But how do you maintain the skill level while roughly doubling in size each year? We rely on the Lake Wobegon Strategy, which says only hire candidates who are above the mean of your current employees. An alternative strategy (popular in the dot-com boom period) is to justify a hire by saying "this candidate is clearly better than at least one of our current employees." The following graph compares the mean employee skill level of two strategies: hire-above-the-mean (or Lake Wobegon) in blue and hire-above-the-min in red. I ran a simulation of 1000 candidates with skill level sampled uniformly from the 0 to 100th percentile (but evaluated by the interview process with noise of ±15%) starting from a core team of 10 employees with mean 75 and min 65. You can see how hire-above-the-min leads to a precipitous drop in skill level; one we've been able to avoid.
The hiring problem

• A small startup is trying to grow into a world-dominating company
• What should its hiring strategy be?
  – Interviews applicants
    • No control over quality
  – Wants to hire the best applicants
  – But also wants to grow
    • Needs working bodies now
The (basic) secretary problem

Applicants for a position are interviewed in random order
The (basic) secretary problem

• Applicants for a position are interviewed in random order
• After each interview, learn the relative rank of the applicant compared to the previous
• Must decide whether to hire or reject immediately
  – All decisions final
• Goal: Maximize probability of getting best
Importance of the problem

• Captures the problem of choosing the best of sequentially presented alternatives
  – Fundamental issue: irrevocable decision making in the presence of uncertainty

• Amenable to many variations/extensions
  – Dozens of papers on different versions, more on applications, etc

• Adaptable to a variety of settings
  – Simplicity and generality of the problem are a feature
Hiring vs secretary problem

The hiring problem focuses on the tradeoff between the rate at which the company grows and the quality of employees

- Avoids focus on “best”
- No fixed number of employees
- Must make sense as a long-term plan for growth
- Online bicriteria optimization
- Score vs rank
Lake Wobegon strategies

• Hire any candidate that is above the average employee at the time of the interview
• Lake Wobegon: fictional town where
  “…all the women are strong, all the men are good-looking, and all the children are above average…”
• Every employee is indeed better than the average when hired!
Lake Wobegon strategies

• Seem intuitively natural and justifiable
  – It would make sense to HR

• Assumption: objective notion of quality
  – IQ/aptitude tests?

• Different notions of “average”
  – We examine hiring above mean and median
Lake Wobegon strategies

Leads to a world of continual improvement, with humorous implications

– Explains why everyone thinks they are better than their boss
– Grad students > Assistant profs > Associate profs > Tenured profs > Retired profs …
The model

• Interview applicants sequentially
• During interview of $i$-th applicant:
  – Observe quality score $Q(i)$
    • Assume iid
  – Immediately decide whether to hire or reject
    • Hired employees take the job
Uniformity assumption

- We assume $Q(i)$ are iid Unif(0,1)
- For any other distribution eg, N(0,1)
  
  $F(x) = \Pr(X \leq x) \leftrightarrow F^{-1}(x) \sim \text{Unif}(0,1)$

- Think of the $Q(i)$ as percentiles of distribution
Uniformity assumption

- We assume $Q(i)$ are iid Unif(0,1)
- For any other distribution eg, N(0,1)
  \[ F(x) = \Pr(X \leq x) \iff F^{-1}(x) \sim \text{Unif}(0,1) \]
- Think of the $Q(i)$ as percentiles of distribution

\[ Q(i) \sim \text{Unif}(0,1) \]
\[ Q(i) \approx 0.84 \]
Two natural strategies

- Hire above fixed threshold
- Hire above the best current employee
Hire above a threshold

• Suppose we hire all applicants whose quality scores lie above some threshold \( t \)
• For small \( t \), basically hiring everyone
• For large \( t \), all employees good, but
  – No continual improvement as company grows
  – Hiring rate 1 − \( t \) : particularly bad when the company is small
  – Stark tradeoff between quality and growth rate
Hire above the max

• Suppose we hire all applicants whose quality scores lie above the current best employee
• Guarantees great employees …
• But hiring is much too slow
• Analysis useful for our other schemes
Hire above the max

• Start with 1 employee with quality $q = 1 - g$
• Let $Q(n)$ be the quality of the $n$th hire
• Define $\text{gap} \ G(n) = 1 - Q(n)$ ($G(0) = g$)
  – Analysis is easier thinking of gap to 1
    $G(n) \sim \text{Unif} \ (0, G(n-1))$
    $E[G(n) | G(n-1)] = G(n-1) / 2$
    $E[G(n)] = g / 2^n$
Lognormal distributions

\[ G(n) \sim \text{Unif}(0, G(n-1)) \sim G(n-1) \text{Unif}(0,1) \]

\[ G(n) = g \prod_{i=1}^{n} U(i) \quad U(i) \sim \text{Unif}(0,1) \]

\[ \ln G(n) = \ln g + \sum_{i=1}^{n} \ln U(i) \]

\[ \approx \ln g + N(-n, n) \]

\[ = N(\ln [g/e^n], n) \]

\( G(n) \) is lognormal (approximately)
Number of interviews

• Start with one employee with gap \( g \)
• First hire has gap \( G(1) \sim g \text{ Unif}(0,1) \)

\[
\mathbb{E}[\# \text{ interviews until 2nd hire}]
\]

\[
= \mathbb{E} \left[ \text{Geometric}(G(1)) \right]
\]

\[
= \mathbb{E} \left[ \frac{1}{G(1)} \right] = \mathbb{E} \left[ \frac{1}{g \text{Unif}(0,1)} \right] = \int_0^1 \frac{1}{gx} \, dx = \infty
\]

• Not good!
Hire above the max: verdict

• Recap: \[ \ln G(n) \approx N \left( \ln \left[ g/e^n \right], n \right) \]
  \[ \mathbb{E}[G(n)] = g/2^n \]
  
• Multiplicative behavior of gap leads to (approximately) lognormal distribution

• Highly skewed
  – Median is \( g/e^n \), but mean is \( g/2^n \)

• Expected number of interviews becomes infinite

• Conclusion: Not a practical strategy
Summary so far

• Threshold strategy
  – No continual improvement
  – Very slow initial hiring
    • But scales nicely: constant interviews/hire

• Max strategy
  – Achieves high quality with few hires
  – Very, very slow hiring

• Looking for a better balance …
Lake Wobegon results

• For hiring above mean/median
  – Expected quality as company grows
  – Expected # of interviews to grow company
  – Comparison

• Weak concentration results
  – First few hires very important
  – As company grows, quality distribution has lognormal characteristics
Hiring above the mean

• Start with 1 employee with quality $q = 1 - g$
• Hire applicant if quality score exceeds mean of current employees’ scores
• Let
  – $A(n)$ denote mean employee quality after $n$ hires
  – $G(n) = 1 - A(n)$ denote the corresponding gap
Representaion lemma

• Quality of \((n+1)\)st hire is \(\text{Unif}(A(n),1)\)
• Corresponding gap is \(\text{Unif}(0,G(n))\)

\[
G(n + 1) = \frac{n + 1}{n + 2} G(n) + \frac{1}{n + 2} \text{Unif}(0, G(n))
\]

\[
\sim G(n) \left(1 - \frac{\text{Unif}(0, 1)}{n + 2}\right)
\]
Natural generalization

The conditional distribution of $G(n + t)$ given $G(n)$ is

$$G(n + t) \mid G(n)$$

$$\sim G(n) \prod_{i=1}^{t} \left(1 - \frac{U(i)}{n + i + 1}\right)$$

where the $U(i)$ are iid Unif(0,1)
Analysis of expectation

• Since \( G(n) \sim g \prod_{i=1}^{n} \left( 1 - \frac{U(i)}{i + 1} \right) \)

\[
\mathbb{E}[G(n)] = g \prod_{i=1}^{n} \left( 1 - \frac{1}{2(i+1)} \right) \approx g \sum_{i=1}^{n} e^{-1/(2(i+1))} = \Theta(1/\sqrt{n})
\]

• Compare to other strategies
  – Hire above threshold: \( \Theta(1) \)
  – Hire above max: \( \Theta(1/2^n) \)
Number of interviews

• Distribution of interviews until next hire is Geometric($G(n)$)

• Let $T(n)$ be # of interviews until $n$ hires

$$\mathbb{E}[T(n)] = \sum_{i=0}^{n-1} \mathbb{E} \left[ \frac{1}{G(i)} \right]$$

• Compute $\mathbb{E}[1/G(i)]$ explicitly via representation lemma

$$\mathbb{E}[T(n)] = \sum_{i=0}^{n-1} \mathbb{E} \left[ \frac{1}{G(i)} \right] = \sum_{i=0}^{n-1} \Theta \left( \sqrt{i} \right) = \Theta \left( n^{3/2} \right)$$

• Compare to other strategies
  – Hire above threshold: $\Theta(n)$
  – Hire above max: $\infty$
Convergence and concentration

• Recall

\[ G(n + t) \mid G(n) \sim G(n) \prod_{i=1}^{t} \left( 1 - \frac{U(i)}{n + i + 1} \right) \]

\[ \ln G(n) = \ln g + \sum_{i=1}^{n} \ln(1 - U(i) / (i + 1)) \]

• Intuition: sum of independent random variables should give lognormal, but because not iid, it is not clear
  – First few hires have bigger effect than later hires
Convergence and concentration

\[ Y(n, m) \triangleq \ln \frac{G(n)}{G(m)} \sim \sum_{i=1}^{n-m} \ln \left( 1 - \frac{U(i)}{i + m + 1} \right) \]

\[ Z(n, m) = Y(n, m) - \mathbf{E}[Y(n, m)] \]

- \( Z(n,m) \) is a sum of independent random variables with mean 0 and vanishing variance
- As \( n \to \infty \) with \( m \) fixed, \( Z(n,m) \) converges
- As \( m = cn \to \infty \), \( Z(n,m) \) looks Gaussian
Convergence and concentration

• Translation:
  – After “first few” hires, $G(n)$ looks lognormal
    • First few is more than constant number
  – Tight concentration results not possible
  – Can still get weak concentration through martingales
    • Instead of asymptotic limiting statements
Hiring above the median

• Start with 1 employee with quality $q = 1 - g$

• Model: when we have $2k + 1$ employees
  – Let $M(k) = \text{median of the scores}$
  – Hire next 2 applicants with scores above $M(k)$
  – Determines $M(k + 1)$

• Define $G^*(k) = 1 - M(k)$
  – $G^*(k)$ is gap when there are $2k + 1$ employees
Representation lemma

Condition on ranking of first $2k + 1$ employees

$M(k)$

```
[Red] [Red] [Red] [Red] [White] [Blue] [Blue] [Blue] [Blue]
```
Representation lemma

Condition on ranking of first $2k + 1$ employees

Each score $\sim \text{Unif}(M(k), 1)$
Representation lemma

Condition on ranking of first $2k + 1$ employees

Each score $\sim \text{Unif}(M(k), 1)$

2 new hires
Each score $\sim \text{Unif}(M(k), 1)$
Representation lemma

Condition on ranking of first $2k + 1$ employees

Each score $\sim \text{Unif}(M(k), 1)$

$k + 2$ scores

Quality
Representation lemma

Condition on ranking of first $2k + 1$ employees

$M(k) \quad M(k+1)$

Each score $\sim M(k)\text{Unif}(0,1)$

$k + 2$ scores

Quality
Representation lemma

Condition on ranking of first $2k + 1$ employees

$M(k) \quad M(k+1)$

$M(k+1) = M(k)(\text{min of } k + 2 \text{ Unif}(0,1) \text{ scores})$
Representation lemma

Condition on ranking of first $2k + 1$ employees

$M(k) \quad M(k+1)$

$M(k+1) = M(k)(\text{Beta}(k+2, 1))$
Representation lemma

Condition on ranking of first $2k + 1$ employees

$$M(k) \quad M(k+1)$$

$$G^*(k + 1) \mid G^*(k)$$

$$\sim G^*(k)\text{Beta}(k + 2, 1)$$

Each score ~ Unif$(M(k), 1)$

$k + 2$ scores

Quality
Comparison

• Recall for hiring above the mean

\[ G(n) \sim g \prod_{i=1}^{n} \left(1 - \frac{U(i)}{i + 1}\right) \]

• Now for hiring above the median

\[ G^*(k) \sim g \prod_{i=1}^{k} \text{Beta}(i + 1, 1) \]

• Analysis for median analogous to mean
Comparison

• For hiring above the mean

\[ \mathbb{E}[G(n)] = \Theta \left( \frac{1}{\sqrt{n}} \right) \quad \mathbb{E}[T(n)] = \Theta \left( n^{3/2} \right) \]

• For hiring above the median

\[ \mathbb{E}[G^*(k)] = \Theta \left( \frac{1}{k} \right) \quad \mathbb{E}[T^*(k)] = \Theta \left( k^2 \right) \]

• Similar convergence/concentration results
Caveat: Apples vs Microsofts

- Hiring above median leads to higher quality but slower growth, than hiring above mean
- Wait: We analyzed
  - Mean quality for hiring above mean
  - Median quality for hiring above median
- Can show mean quality of hiring above median is $\Theta((\log k)/k)$
- We don’t know how to argue the median quality score of hiring above mean
Variations and extensions

- Interview preprocessing
  - Quality distribution may change with time
  - “Weed out” bad candidates before interview

- Use different quantiles
  - Median is “hire two, move up one”
  - Can do “hire A, move up B”
  - Will amount to general beta distributions

- Errors
  - Only get noisy estimates of true quality
  - What is quality anyway?
Hiring and firing

• Firing allows us to remove low performers as the company grows
• Gives further optimization of the number of employees and their average quality
• But introduces dependencies that make analysis difficult
Rank and yank firing

• Periodically rank all employees
• Fire the bottom 10% (or any fraction $f$)
  – Regardless of how good the employees are!
  – Jack Welch did this at GE for a long time
Rank and yank firing

- Periodically rank all employees
- Fire the bottom 10% (or any fraction $f$)
- Easily modeled with hiring above the median

Median $M$
Rank and yank firing

- Periodically rank all employees.
- Fire the bottom 10% (or any fraction $f$).
- Easily modeled with hiring above the median

Median $M$

Each score $\sim \text{Unif}(M,1)$
Rank and yank firing

- Periodically rank all employees
- Fire the bottom 10% (or any fraction $f$)
- Easily modeled with hiring above the median

\[
\text{Median } M
\]

Each Score $\sim \text{Unif}(M,1)$
Rank and yank firing

- Periodically rank all employees
- Fire the bottom 10% (or any fraction $f$)
- Easily modeled with hiring above the median

New median $M^*$

Yields a representation lemma

Each score $\sim \text{Unif}(M,1)$
Summary

• Introduced hiring problem
  – Natural analog of secretary problem
  – Useful paradigm for decision making under uncertainty

• Modeled/analyzed natural strategies
  – Many interesting probabilistic phenomena
  – Lots of potential for extensions and variations
Open questions

• Differences between strategies that use score and those that only use ranking?
• Lower bounds relating (expected) quality and (expected) time to hire?
• Optimality/near-optimality of these strategies?
• Strategies giving all possible time/quality tradeoffs?
• Expanding/elaborating on interesting variations
Another hiring strategy we use is no hiring manager. Whenever you give project managers responsibility for hiring for their own projects they’ll take the best candidate in the pool, even if that candidate is sub-standard for the company, because every manager wants some help for their project rather than no help. That’s why we do all hiring at the company level, not the project level. First we decide which candidates are above the hiring threshold, and then we decide what projects they can best contribute to. The orange line in the graph above is a simulation of the hiring-manager strategy, with the same candidates and the same number of hires as the no-hiring-manager strategy in blue. Employees are grouped into pools of random size from 2 to 14 and the hiring manager chooses the best one. We’re pleased that these little simulations show our hiring strategy is on top. You can learn more about our hiring and working philosophy.
Xie xie!

ravikumar@yahoo-inc.com