Adventures in Crowdsourcing

Panos Ipeirotis

Thanks to: Jing Wang, Marios Kokkodis, Foster Provost, Josh Attenberg, Victor Sheng, Evgeniy Gabrilovich, Chun How Tan, Ari Kobren, Gabrielle Paolaci, Jesse Chandler

Twitter: @ipeirotis

“A Computer Scientist in a Business School”
http://behind-the-enemy-lines.com
Integrate machine and human intelligence
Create hybrid “intelligence integration” processes
With *paid* users and with *unpaid* users
Application
Detect Inappropriate content

- Need to detect inappropriate content
  - Ad placement, FB feed, links in forums, etc
- Ad hoc topics, with no existing training data
  - Hate speech, Violence, Guns & Bombs, Gossip…
- Classification models needed within days
- **Crowdsourcing** allows for fast data collection
  - using Mechanical Turk, oDesk, etc
  - labor is accessible on demand
  - but quality may be lower than experts
### Amazon Mechanical Turk

#### All HITs
1-10 of 1984 Results

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Reward</th>
<th>Time Allotted</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sam GONZALES</td>
<td>Dec 13, 2010 (1 week 2 days)</td>
<td>$0.01</td>
<td>30 minutes</td>
<td>39172</td>
</tr>
<tr>
<td>Chris Callison-Burch</td>
<td>Dec 31, 2010 (3 weeks 6 days)</td>
<td>$0.05</td>
<td>15 minutes</td>
<td>14240</td>
</tr>
<tr>
<td>nutella42</td>
<td>Dec 17, 2010 (2 weeks)</td>
<td>$0.08</td>
<td>30 minutes</td>
<td>2446</td>
</tr>
<tr>
<td>Jaime Arquello</td>
<td>Dec 10, 2010 (7 days)</td>
<td>$0.03</td>
<td>5 minutes</td>
<td>1952</td>
</tr>
<tr>
<td>Andy K</td>
<td>Dec 9, 2010 (6 days 2 hours)</td>
<td>$0.03</td>
<td>60 minutes</td>
<td>1949</td>
</tr>
</tbody>
</table>

#### Identify Arabic Dialect in Text
Requester: Chris Callison-Burch

**HIT Expiration Date:** Dec 31, 2010 (3 weeks 6 days)
**Reward:** $0.05
**Time Allotted:** 15 minutes
**HITs Available:** 14240

#### POI Verification for USA Cities
Requester: nutella42

**HIT Expiration Date:** Dec 17, 2010 (2 weeks)
**Reward:** $0.08
**Time Allotted:** 30 minutes
**HITs Available:** 2446

#### Preference Judgements between Search Engine Results
Requester: Jaime Arquello

**HIT Expiration Date:** Dec 10, 2010 (7 days)
**Reward:** $0.03
**Time Allotted:** 5 minutes
**HITs Available:** 1952

#### Keyword Category Verification
Requester: Andy K

**HIT Expiration Date:** Dec 9, 2010 (6 days 2 hours)
**Reward:** $0.03
**Time Allotted:** 60 minutes
**HITs Available:** 1949
Help Classify Arabic into Dialects!

This task is for Arabic speakers who understand the different local Arabic dialects (اللهجات المحلية، أو التأريخ)، and can distinguish them from Fusha Arabic (العربية الفصحى).

Below, you will see several Arabic sentences. For each sentence:

1. Tell us how much dialect (عربية) is in the sentence, and then
2. Tell us which Arabic dialect the writer intends.

This following map explains the dialects:

<table>
<thead>
<tr>
<th>Which Dialect?</th>
<th>Dialect Level</th>
<th>Sentence</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose level first</td>
<td>Choose level...</td>
<td>#1 خليه براحته يا جماعة الخير.</td>
<td></td>
</tr>
<tr>
<td>Choose level first</td>
<td>Choose level...</td>
<td>#2 اللهم يا هلال</td>
<td></td>
</tr>
<tr>
<td>Choose level first</td>
<td>Choose level...</td>
<td>#3 سبحان الله وتعجب</td>
<td></td>
</tr>
</tbody>
</table>
Example: Build an “Adult Content” Classifier

- Need a large number of labeled sites for training
- Get people to look at sites and label them as:
  - **G** (general audience)
  - **PG** (parental guidance)
  - **R** (restricted)
  - **X** (porn)

Cost/Speed Statistics

- **Undergrad intern**: 200 websites/hr, cost: $15/hr
- **Mechanical Turk**: 2500 websites/hr, cost: $12/hr
Bad news: Spammers!

Worker ATAMRO447HWJQ labeled X (porn) sites as G (general audience)
Challenges

- We do not know the true category for the objects
  - Available only after (costly) manual inspection
- We do not know quality of the workers
- We want to label objects with true categories
- We want (need?) to know the quality of the workers
Redundant votes, infer quality

Look at our lazy friend ATAMRO447HWJQ together with other 9 workers

Using redundancy, we can compute error rates for each worker
Expectation Maximization Estimation

Iterative process to estimate worker error rates

1. **Initialize** “correct” label for each object (e.g., use majority vote)
2. **Estimate** error rates for workers (using “correct” labels)
3. **Estimate** “correct” labels (using error rates, weight worker votes according to quality)
4. Go to Step 2 and iterate until convergence

Error rates for ATAMRO447HWJQ
- $P[G \rightarrow G]=99.947\%$
- $P[G \rightarrow X]=0.053\%$
- $P[X \rightarrow G]=99.153\%$
- $P[X \rightarrow X]=0.847\%$

Our friend ATAMRO447HWJQ marked **almost all** sites as $G$. Clickety clickey clo…
**Challenge: Humans are biased!**

<table>
<thead>
<tr>
<th>Error rates for the CEO, providing “expert” labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P[G \rightarrow G] = 20.0%$</td>
</tr>
<tr>
<td>$P[P \rightarrow G] = 0.0%$</td>
</tr>
<tr>
<td>$P[R \rightarrow G] = 0.0%$</td>
</tr>
<tr>
<td>$P[X \rightarrow G] = 0.0%$</td>
</tr>
</tbody>
</table>

- We have **85% G sites, 5% P sites, 5% R sites, 5% X sites**
- Error rate of **spammer** (all G) = $0\% \times 85\% + 100\% \times 15\% = 15\%$
- Error rate of **biased worker** = $80\% \times 85\% + 100\% \times 5\% = 73\%$

**False positives: Legitimate workers appear to be spammers**
(important note: bias is not just a matter of “ordered” classes)
Solution: Fix bias first, compute error rate afterwards

Error Rates for CEO of AdSafe

<table>
<thead>
<tr>
<th>Event</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>P[G → G]</td>
<td>20.0%</td>
</tr>
<tr>
<td>P[P → G]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[R → G]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[X → G]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[G → P]</td>
<td>80.0%</td>
</tr>
<tr>
<td>P[P → P]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[R → P]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[X → P]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[G → R]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[P → R]</td>
<td>100.0%</td>
</tr>
<tr>
<td>P[R → R]</td>
<td>100.0%</td>
</tr>
<tr>
<td>P[X → R]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[G → X]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[P → X]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[R → X]</td>
<td>0.0%</td>
</tr>
<tr>
<td>P[X → X]</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

- When biased worker says G, it is **100% G**
- When biased worker says P, it is **100% G**
- When biased worker says R, it is **50% P, 50% R**
- When biased worker says X, it is **100% X**

Small ambiguity for “R-rated” votes but other than that, fine!
Question: How to pay workers?

- Naïve solution: Have a quality-score threshold
- **Threshold-ing rewards gives wrong incentives:**
  - Decent (but still useful) workers get fired
  - Uncertainty near the decision threshold
Quality-sensitive Payment

- Set quality goal and price (e.g., $1 for 90%)
  - For workers above goal: Pay full price
  - For others: Payment divided with redundancy needed to reach goal
    - Need 3 workers with 80% accuracy $\Rightarrow$ Payment = $0.33$
    - Need 9 workers with 70% accuracy $\Rightarrow$ Payment = $0.11$
Instead of blocking: Quality-sensitive Payment

- Estimate payment level based on quality
  - Set acceptable quality (e.g., 99% accuracy)
  - For workers above quality specs: Pay full price
  - For others: Estimate level of redundancy to reach acceptable quality (e.g., Need 5 workers with 90% accuracy or 13 workers with 80% accuracy to reach 99% accuracy;)
  - Pay full price divided by level of necessary redundancy
- **Uncertainty** penalty: **Pay less** for uncertain estimates (for workers with short working histories)
- **Refund** underpayment when quality estimate more certain
Real-Time Payment and Reimbursement

Example of the piece-rate payment of a worker

<table>
<thead>
<tr>
<th>#Tasks</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate Payment (cents)</td>
<td>11</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>40</td>
</tr>
</tbody>
</table>

Fair Payment
Real-Time Payment and Reimbursement

Example of the piece-rate payment of a worker

<table>
<thead>
<tr>
<th>#Tasks</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate Payment (cents)</td>
<td>11</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>40</td>
</tr>
</tbody>
</table>

Fair Payment: 40

Piece-rate Payment

Potential "Bonus"

Payment

Number of Tasks
## Real-Time Payment and Reimbursement

Example of the piece-rate payment of a worker

<table>
<thead>
<tr>
<th>#Tasks</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate Payment (cents)</td>
<td>11</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>40</td>
</tr>
</tbody>
</table>

**Fair Payment: 40**

- **Potential “Bonus”**
- **Reimbursement**
- **Payment**

Number of Tasks

10

20
Real-Time Payment and Reimbursement

Example of the piece-rate payment of a worker

<table>
<thead>
<tr>
<th>#Tasks</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate Payment (cents)</td>
<td>11</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>40</td>
</tr>
</tbody>
</table>

Fair Payment: 40
Real-Time Payment and Reimbursement

Example of the piece-rate payment of a worker

<table>
<thead>
<tr>
<th>#Tasks</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Infinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate Payment (cents)</td>
<td>11</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>40</td>
</tr>
</tbody>
</table>

Fair Payment: 40
Improving worker participation

- With just labeling, workers are passively labeling the data that we give them.

- But this can be wasteful when positive cases are sparse.

- Why not asking the workers to search themselves and find training data?
Guided Learning

Ask workers to find example web pages (great for “sparse” content)

After collecting enough examples, easy to build and test web page classifier

http://url-collector.appspot.com/allTopics.jsp
Limits of Guided Learning

- No incentives for workers to find “new” content
- After a while, submitted web pages similar to already submitted ones
- No improvement for classifier
The result? Blissful ignorance…

- Classifier *seems* great: Cross-validation tests show excellent performance

- Alas, classifier fails: The “*unknown unknowns*” ™
  
  No similar training data in training set
  
  “*Unknown unknowns*” = classifier fails with high confidence
Beat the Machine!

Ask humans to find URLs that

- *the classifier will classify incorrectly*
- *another human will classify correctly*

Example:

*Find hate speech pages that the machine will classify as benign*

http://adsafe-beatthemachine.appspot.com/
Beat the Machine!

Incentive structure:

- **$1 if you “beat the machine”**
- **$0.001 if the machine already knows**

Example:

*Find hate speech pages that the machine will classify as benign*

http://adsafe-beatthemachine.appspot.com/
Error rate for probes significantly higher than error rate on (stratified) random data (10x to 100x higher than base error rate)

<table>
<thead>
<tr>
<th>#</th>
<th>Category</th>
<th>Tasks Running</th>
<th>URL's gathered</th>
<th>Correct URL's gathered</th>
<th>Total Bonus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identify pages that contain hate speech on the web (hat)</td>
<td>206</td>
<td>1023</td>
<td>161</td>
<td>75516</td>
</tr>
<tr>
<td>2</td>
<td>Identify pages related to illegal drug use on the web (drg)</td>
<td>100</td>
<td>500</td>
<td>26</td>
<td>9114</td>
</tr>
<tr>
<td>3</td>
<td>Identify pages that contain reference to alcohol (alc)</td>
<td>100</td>
<td>475</td>
<td>144</td>
<td>55149</td>
</tr>
<tr>
<td>4</td>
<td>Identify adult-related pages (adt)</td>
<td>174</td>
<td>859</td>
<td>132</td>
<td>63523</td>
</tr>
</tbody>
</table>
No money?

- What if we want to engage users without paying them?
Google Knowledge Graph

"Things not Strings"
Still incomplete…

- “Date of birth of Bayes” (…uncertain…)
- “Symptom of strep throat”
- “Side effects of treximet”
- “Who is Cristiano Ronaldo dating”
- “When is Jay Z playing in New York”
- “What is the customer service number for Google”
- …
The Google mission…

We have a billion users…
Leverage their knowledge…

“Let’s create a new crowdsourcing system…”
Ideally…
But often...
The common solution...
“Crowdsourcing in a **predictable** manner, with **knowledgeable** users, without introducing **monetary rewards**”
What is a symptom of **Morgellons**

- Red eye
- Choreoathetosis
- Skin lesion
- Insomnia
- I don't know

Correct Answers: 33/67  Correct (%): 49%

Question 1 out of 10
Calibration vs. Collection

- **Calibration** questions (known answer): Evaluating user competence on topic at hand
- **Collection** questions (unknown answer): Asking questions for things we do not know
- *Trust more answers coming from competent users*
Challenges

- Why would anyone come and play this game?
- Why would knowledgeable users come?
- Wouldn’t it be simpler to just pay?
Quiz on disease symptoms
Test how well you can recognize various disease symptoms
www.quizz.us
Treat Quizz as eCommerce Site

Feedback:
Value of user
Example of Targeting: Medical Quizzes

- Our initial goal was to use medical topics as a evidence that some topics are not crowdsourcable

- Our hypothesis failed: They were the best performing quizzes…

- Users coming from sites such as Mayo Clinic, WebMD, … (i.e., “pronsumers”, not professionals)
Immediate feedback helps

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show if user answer correct</td>
<td>+2.4%</td>
</tr>
<tr>
<td>Show the correct answer</td>
<td>+20.4%</td>
</tr>
<tr>
<td>Score: % of correct answers</td>
<td>+2.3%</td>
</tr>
<tr>
<td>Score: # of correct answers</td>
<td>-2.2%</td>
</tr>
<tr>
<td>Score: Information gain</td>
<td>+4.0%</td>
</tr>
<tr>
<td>Show statistics for performance of other users</td>
<td>+9.8%</td>
</tr>
<tr>
<td>Leaderboard based on percent correct</td>
<td>-4.8%</td>
</tr>
<tr>
<td>Leaderboard based on total correct answers</td>
<td>-1.5%</td>
</tr>
</tbody>
</table>

- Knowing the correct answer 10x more important than knowing whether given answer was correct
- Conjecture: Users also want to learn
### Showing score moderately helpful

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show if user answer correct</td>
<td>+2.4%</td>
</tr>
<tr>
<td>Show the correct answer</td>
<td>+20.4%</td>
</tr>
<tr>
<td><strong>Score: % of correct answers</strong></td>
<td>+2.3%</td>
</tr>
<tr>
<td><strong>Score: # of correct answers</strong></td>
<td>-2.2%</td>
</tr>
<tr>
<td><strong>Score: Information gain</strong></td>
<td>+4.0%</td>
</tr>
<tr>
<td>Show statistics for performance of other users</td>
<td>+9.8%</td>
</tr>
<tr>
<td>Leaderboard based on percent correct</td>
<td>-4.8%</td>
</tr>
<tr>
<td>Leaderboard based on total correct answers</td>
<td>-1.5%</td>
</tr>
</tbody>
</table>

- Be careful what you incentivize 😊
- “Total Correct” incentivizes quantity, not quality
## Competitiveness helps

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show if user answer correct</td>
<td>+2.4%</td>
</tr>
<tr>
<td>Show the correct answer</td>
<td>+20.4%</td>
</tr>
<tr>
<td>Score: % of correct answers</td>
<td>+2.3%</td>
</tr>
<tr>
<td>Score: # of correct answers</td>
<td>-2.2%</td>
</tr>
<tr>
<td>Score: Information gain</td>
<td>+4.0%</td>
</tr>
<tr>
<td><strong>Show statistics for performance of other users</strong></td>
<td><strong>+9.8%</strong></td>
</tr>
<tr>
<td>Leaderboard based on percent correct</td>
<td>-4.8%</td>
</tr>
<tr>
<td>Leaderboard based on total correct answers</td>
<td>-1.5%</td>
</tr>
</tbody>
</table>
Leaderboards are tricky!

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show if user answer correct</td>
<td>+2.4%</td>
</tr>
<tr>
<td>Show the correct answer</td>
<td>+20.4%</td>
</tr>
<tr>
<td>Score: % of correct answers</td>
<td>+2.3%</td>
</tr>
<tr>
<td>Score: # of correct answers</td>
<td>-2.2%</td>
</tr>
<tr>
<td>Score: Information gain</td>
<td>+4.0%</td>
</tr>
<tr>
<td>Show statistics for performance of other users</td>
<td>+9.8%</td>
</tr>
<tr>
<td><strong>Leaderboard based on percent correct</strong></td>
<td><strong>-4.8%</strong></td>
</tr>
<tr>
<td><strong>Leaderboard based on total correct answers</strong></td>
<td><strong>-1.5%</strong></td>
</tr>
</tbody>
</table>

- Initially, strong positive effect
- Over time, effect became strongly negative
- All-time leaderboards considered harmful
Comparison with paid crowdsourcing

Unpaid users vs. hourly (oDesk) vs. piecemeal (MTurk)

% correct

Submitted answers (log)

Source
- MTurk
- oDesk
- Quizz
Citizen Science Applications

- Google gives $10K/month to nonprofits in ad budget

- Climate CoLab experiment running
  - Doubled traffic with only $20/day
  - Targets political activist groups (not only climate)

- Additional experiments: Crowdcrafting, ebird, Weendy
How can I get rid of users?
Your workers behave like my mice!

An unexpected connection…
Your workers behave like my mice!

Eh?
Your workers want to use only their **motor skills**, not their **cognitive skills**.
The Biology Fundamentals

- Brain functions are biologically expensive (20% of total energy consumption in humans)

- Motor skills are more energy efficient than cognitive skills (e.g., walking)

- Brain tends to delegate easy tasks to part of the neural system that handles motor skills
An unexpected connection at the NAS “Frontiers of Science” conf.

Your workers want to use only their **motor skills**, not their **cognitive skills**

*Don Cooper*
Department of Psychology & Neuroscience

*Makes sense*
An unexpected connection at the NAS “Frontiers of Science” conf.

And here is how I train my mice to behave…
The Mice Experiment

Cognitive
Solve maze
Find pellet

Motor
Push lever three times
Pellet drops
How to Train the Mice?

Confuse motor skills! Reward cognition!

Don Cooper
Department of Psychology & Neuroscience

I should try this the moment that I get back to my room
Punishing Worker’s Motor Skills

- **Punish bad answers** with frustration of motor skills (e.g., add delays between tasks)
  - “Loading image, please wait…”
  - “Image did not load, press here to reload”
  - “404 error. Return the HIT and accept again”

→ Make this **probabilistic** to keep feedback implicit
Misery

Posted by danielb on June 22, 2009 at 10:10am

Misery is a module designed to make life difficult for certain users.

It can be used:

- As an alternative to banning or deleting users from a community.
- As a means by which to punish members of your website.
- To delight in the suffering of others.

Currently you can force users (via permissions/roles, editing their user account, or using Troll IP blacklists) to endure the following misery:

- **Delay**: Create a random-length delay, giving the appearance of a slow connection. (by default this happens 40% of the time)
- **White screen**: Present the user with a white-screen. (by default this happens 10% of the time)
- **Wrong page**: Redirect to a random URL in a predefined list. (by default this happens 0% of the time)
- **Random node**: Redirect to a random node accessible by the user. (by default this happens 10% of the time)
- **403 Access Denied**: Present the user with an "Access Denied" error. (by default this happens 10% of the time)
- **404 Not Found**: Present the user with a "Not Found" error. (by default this happens 10% of the time)
Experimental Summary (I)

- Spammer workers quickly abandon
  - No need to display scores, or ban
  - Low quality submissions from ~60% to ~3%
  - Half-life of low-quality from 100+ HITs to less than 5

- Good workers unaffected
  - No significant effect on participation of workers with good performance
  - Lifetime of participants unaffected
  - Longer response times (*after* removing the “intervention delays”; that was puzzling)
Remember, scheme was for *training* the mice…

Indeed, 15%-20% of the spammers start submitting good work!

?????
Two key questions

- Why response time was slower for some good workers?
- Why some low quality workers start working well?
System 1: “Automatic” actions

System 2: “Intelligent” actions
System 1 Tasks

- Detect that one object is more distant than another.
- Orient to the source of a sudden sound.
- Complete the phrase “bread and...”
- Make a “disgust face” when shown a horrible picture.
- Detect hostility in a voice.
- Answer to $2 + 2 = ?$
- Read words on large billboards.
- Drive a car on an empty road.
- Find a strong move in chess (if you are a chess master).
- Understand simple sentences.
System 2 Tasks

- Focus attention on the clowns in the circus.

- Look for a woman with white hair.

- Count the occurrences of the letter a in a page of text.

- Compare two washing machines for overall value.

- Check the validity of a complex logical argument.
Performing Well?

Status: Usage of System 1 ("Automatic")

Not Performing Well?
Disrupt and Engage System 2

Performing Well?
Check if System 1 can Handle, remove System 2 stimuli

Status: Usage of System 2 ("Intelligent")

Not Performing Well?
Hell/Slow ban

Out
Thanks!

Q & A?
Effect of Ad Targeting

Perhaps it is just more users?

- **Control**: Ad campaign with no feedback, all keywords across quizzes
- **Treatment**: Ad campaign with feedback enabled

- **Clicks/visitors**: Same
- **Conversion rate**: 34% vs 13% (~3x more users participated)
- **Number of answers**: 2866 vs 279 (~10x more answers submitted)
- **Total Information Gain**: 7560 bits vs 610 bits (~11.5x more bits)