COMPSCI 514: ALGORITHMS FOR DATA SCIENCE

Andrew McGregor

Lecture 3

Last Class We Covered:

- Markov's inequality: the most fundamental concentration bound.
- Algorithmic applications of Markov's inequality, linearity of expectation, and indicator random variables:
    - Counting collisions to estimate CAPTCHA database size.
    - Counting collisions to understand the runtime of hash tables with random hash functions.

Last Class We Covered:

- Markov's inequality: the most fundamental concentration bound.
- Algorithmic applications of Markov's inequality, linearity of expectation, and indicator random variables:
  - Counting collisions to estimate CAPTCHA database size.
  - Counting collisions to understand the runtime of hash tables with random hash functions.
- Collision counting is closely related to the birthday paradox.

Today:

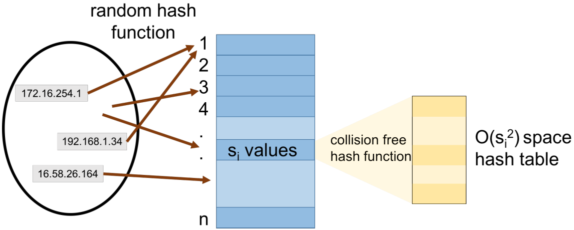
- Finish up random hash functions and hash tables.
- See an application of random hashing to load balancing in distributed systems.
- Through this application learn about:
  - Chebyshev's inequality, which strengthens Markov's inequality.
  - The union bound, for understanding the probabilities of correlated random events.

TWO LEVEL HASHING

Want to preserve $O(1)$ query time while using $O(m)$ space.

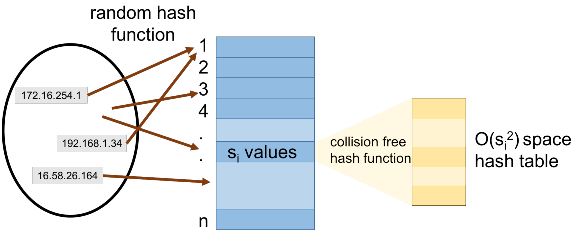
Want to preserve $O(1)$ query time while using $O(m)$ space.

Two-Level Hashing:



Want to preserve $O(1)$ query time while using $O(m)$ space.

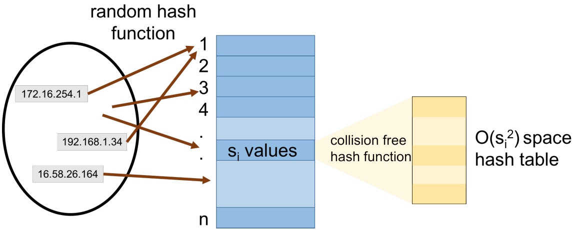
Two-Level Hashing:



- For each bucket with s_i values, pick a collision free hash function mapping $[s_i] \rightarrow [s_i^2]$.

Want to preserve $O(1)$ query time while using $O(m)$ space.

Two-Level Hashing:



- For each bucket with s_i values, pick a collision free hash function mapping $[s_i] \to [s_i^2]$.
- **Just Showed:** A random function is collision free with probability $\geq \frac{7}{8}$ so can just generate a random hash function and check if it is collision free.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions.

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbf{S} = n + \sum_{i=1}^{n} \mathbf{s}_i^2$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$

Collisions again!

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$,

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i] = \frac{1}{n}$.

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i] = \frac{1}{n}$.
- For $j \neq k$,

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \Pr[\mathbf{h}(x_j) = i \cap \mathbf{h}(x_k) = i]$

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

Query time for two level hashing is $O(1)$: requires evaluating two hash functions. What is the expected space usage?

Up to constants, space used is: $\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$

$$\mathbb{E}[\mathbf{s}_i^2] = \mathbb{E}\left[\left(\sum_{j=1}^{m} \mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{j,k\in[m]} \mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right].$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \mathbb{E}\left[\left(\mathbb{I}_{\mathbf{h}(x_j)=i}\right)^2\right] = \Pr[\mathbf{h}(x_j) = i] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \Pr[\mathbf{h}(x_j) = i \cap \mathbf{h}(x_k) = i] = \frac{1}{n^2}$.

x_j, x_k: stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored in hash table at position i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, **h**: random hash function, **S**: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k \in [m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, **h**: random hash function, **S**: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k \in [m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$
$$= \frac{m}{n} + \frac{m(m-1)}{n^2}$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, **h**: random hash function, **S**: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k \in [m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$
$$= \frac{m}{n} + \frac{m(m-1)}{n^2} \leq 2 \text{ (If we set } n = m.)$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

x_j, x_k: stored items, m: # stored items, n: hash table size, **h**: random hash function, **S**: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k \in [m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$
$$= \frac{m}{n} + \frac{m(m-1)}{n^2} \leq 2 \text{ (If we set } n = m.\text{)}$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

Total Expected Space Usage: (if we set $n = m$)

$$\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2]$$

x_j, x_k: stored items, m: # stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

SPACE USAGE

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k\in[m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$

$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$

$$= \frac{m}{n} + \frac{m(m-1)}{n^2} \leq 2 \text{ (If we set } n = m.)$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

Total Expected Space Usage: (if we set $n = m$)

$$\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2] \leq n + n \cdot 2 = 3n = 3m.$$

x_j, x_k: stored items, m: # stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

$$\mathbb{E}[\mathbf{s}_i^2] = \sum_{j,k \in [m]} \mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right]$$
$$= m \cdot \frac{1}{n} + 2 \cdot \binom{m}{2} \cdot \frac{1}{n^2}$$
$$= \frac{m}{n} + \frac{m(m-1)}{n^2} \leq 2 \text{ (If we set } n = m.)$$

- For $j = k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n}$.
- For $j \neq k$, $\mathbb{E}\left[\mathbb{I}_{\mathbf{h}(x_j)=i} \cdot \mathbb{I}_{\mathbf{h}(x_k)=i}\right] = \frac{1}{n^2}$.

Total Expected Space Usage: (if we set $n = m$)

$$\mathbb{E}[\mathbf{S}] = n + \sum_{i=1}^{n} \mathbb{E}[\mathbf{s}_i^2] \leq n + n \cdot 2 = 3n = 3m.$$

Near optimal space with $O(1)$ query time!

x_j, x_k: stored items, m: # stored items, n: hash table size, $\mathbf{h}$: random hash function, $\mathbf{S}$: space usage of two level hashing, $\mathbf{s}_i$: # items stored at pos i.

EFFICIENTLY COMPUTABLE HASH FUNCTION

So Far: we have assumed a **fully random hash function** $\mathbf{h}(x)$ with $\Pr[\mathbf{h}(x) = i] = \frac{1}{n}$ for $i \in 1, \ldots, n$ and $\mathbf{h}(x), \mathbf{h}(y)$ independent for $x \neq y$.

So Far: we have assumed a **fully random hash function** $\mathbf{h}(x)$ with $\Pr[\mathbf{h}(x) = i] = \frac{1}{n}$ for $i \in 1, \ldots, n$ and $\mathbf{h}(x), \mathbf{h}(y)$ independent for $x \neq y$.

- To compute a random hash function we have to store a table of x values and their hash values. Would take at least $O(m)$ space and $O(m)$ query time if we hash m values. Making our whole quest for $O(1)$ query time pointless!

x	**h(x)**
x_1	45
x_2	1004
x_3	10
⋮	⋮
x_m	12

EFFICIENTLY COMPUTABLE HASH FUNCTIONS

What properties did we use of the randomly chosen hash function?

What properties did we use of the randomly chosen hash function?

2-Universal Hash Function (low collision probability). A random hash function from $\mathbf{h} : U \rightarrow [n]$ is two universal if:

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] \leq \frac{1}{n}.$$

What properties did we use of the randomly chosen hash function?

2-Universal Hash Function (low collision probability). A random hash function from $\mathbf{h} : U \rightarrow [n]$ is two universal if:

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] \leq \frac{1}{n}.$$

Exercise: Rework the two level hashing proof to show that this property is really all that is needed.

What properties did we use of the randomly chosen hash function?

2-Universal Hash Function (low collision probability). A random hash function from $\mathbf{h} : U \to [n]$ is two universal if:

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] \leq \frac{1}{n}.$$

Exercise: Rework the two level hashing proof to show that this property is really all that is needed.

When $\mathbf{h}(x)$ and $\mathbf{h}(y)$ are chosen independently at random from $[n]$, $\Pr[\mathbf{h}(x) = \mathbf{h}(y)] = \frac{1}{n}$ (so a fully random hash function is 2-universal)

What properties did we use of the randomly chosen hash function?

2-Universal Hash Function (low collision probability). A random hash function from $\mathbf{h}: U \to [n]$ is two universal if:

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] \leq \frac{1}{n}.$$

Exercise: Rework the two level hashing proof to show that this property is really all that is needed.

When $\mathbf{h}(x)$ and $\mathbf{h}(y)$ are chosen independently at random from $[n]$, $\Pr[\mathbf{h}(x) = \mathbf{h}(y)] = \frac{1}{n}$ (so a fully random hash function is 2-universal)

Efficient Alternative: Let p be a prime with $p \geq |U|$. Choose random $\mathbf{a}, \mathbf{b} \in [p]$ with $\mathbf{a} \neq 0$. Let:

$$\mathbf{h}(x) = (\mathbf{a}x + \mathbf{b} \mod p) \mod n.$$

Another common requirement for a hash function:

Another common requirement for a hash function:

Pairwise Independent Hash Function. A random hash function from $\mathbf{h} : U \to [n]$ is pairwise independent if for all $i, j \in [n]$:

$$\Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = j] = \frac{1}{n^2}.$$

Another common requirement for a hash function:

Pairwise Independent Hash Function. A random hash function from $\mathbf{h} : U \to [n]$ is pairwise independent if for all $i, j \in [n]$:

$$\Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = j] = \frac{1}{n^2}.$$

Breakout: Which is a more stringent requirement? 2-universal or pairwise independent?

Another common requirement for a hash function:

Pairwise Independent Hash Function. A random hash function from $\mathbf{h} : U \to [n]$ is pairwise independent if for all $i, j \in [n]$:

$$\Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = j] = \frac{1}{n^2}.$$

Breakout: Which is a more stringent requirement? 2-universal or pairwise independent?

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] = \sum_{i=1}^{n} \Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = i] = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

Another common requirement for a hash function:

Pairwise Independent Hash Function. A random hash function from $\mathbf{h} : U \to [n]$ is pairwise independent if for all $i, j \in [n]$:

$$\Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = j] = \frac{1}{n^2}.$$

Breakout: Which is a more stringent requirement? 2-universal or pairwise independent?

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] = \sum_{i=1}^{n} \Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = i] = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

A closely related $(\mathbf{a}x + \mathbf{b}) \mod p$ construction gives pairwise independence on top of 2-universality.

Another common requirement for a hash function:

Pairwise Independent Hash Function. A random hash function from $\mathbf{h} : U \to [n]$ is pairwise independent if for all $i, j \in [n]$:

$$\Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = j] = \frac{1}{n^2}.$$

Breakout: Which is a more stringent requirement? 2-universal or pairwise independent?

$$\Pr[\mathbf{h}(x) = \mathbf{h}(y)] = \sum_{i=1}^{n} \Pr[\mathbf{h}(x) = i \cap \mathbf{h}(y) = i] = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

A closely related $(\mathbf{a}x + \mathbf{b}) \mod p$ construction gives pairwise independence on top of 2-universality.

Remember: A fully random hash function is both 2-universal and pairwise independent. But it is not efficiently implementable.

NEXT STEP

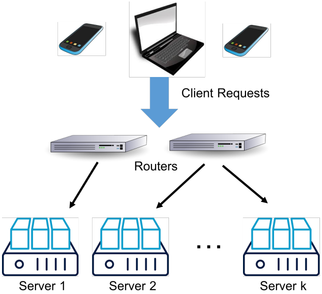
1. We'll consider an application where our toolkit of linearity of expectation + Markov's inequality doesn't give much.

NEXT STEP

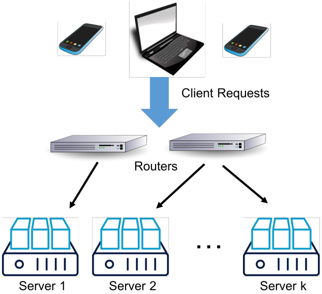
1. We'll consider an application where our toolkit of linearity of expectation + Markov's inequality doesn't give much.

2. Then we'll show how a simple twist on Markov's can give a much stronger result.

Randomized Load Balancing:

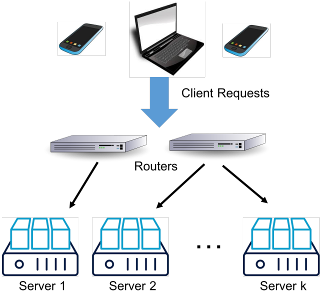


Randomized Load Balancing:



Simple Model: n requests randomly assigned to k servers. How many requests must each server handle?

Randomized Load Balancing:



Simple Model: n requests randomly assigned to k servers. How many requests must each server handle?

Expected Number of requests assigned to server i:

$$\mathbb{E}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathbb{E}[\mathbb{I}_{\text{request } j \text{ assigned to } i}] = \sum_{j=1}^{n} \Pr\left[j \text{ assigned to } i\right] = \frac{n}{k}.$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

Expected Number of requests assigned to server i:

$$\mathbb{E}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathbb{E}[\mathbb{I}_{\text{request } j \text{ assigned to } i}] = \sum_{j=1}^{n} \Pr\left[j \text{ assigned to } i\right] = \frac{n}{k}.$$

If we provision each server be able to handle twice the expected load, what is the probability that a server is overloaded?

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

Expected Number of requests assigned to server i:

$$\mathbb{E}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathbb{E}[\mathbb{I}_{\text{request } j \text{ assigned to } i}] = \sum_{j=1}^{n} \Pr\left[j \text{ assigned to } i\right] = \frac{n}{k}.$$

If we provision each server be able to handle twice the expected load, what is the probability that a server is overloaded?

Applying Markov's Inequality

$$\Pr\left[\mathbf{R}_i \geq 2\mathbb{E}[\mathbf{R}_i]\right] \leq \frac{\mathbb{E}[\mathbf{R}_i]}{2\mathbb{E}[\mathbf{R}_i]} = \frac{1}{2}.$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

Expected Number of requests assigned to server i:

$$\mathbb{E}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathbb{E}[\mathbb{I}_{\text{request } j \text{ assigned to } i}] = \sum_{j=1}^{n} \Pr[j \text{ assigned to } i] = \frac{n}{k}.$$

If we provision each server be able to handle twice the expected load, what is the probability that a server is overloaded?

Applying Markov's Inequality

$$\Pr[\mathbf{R}_i \geq 2\mathbb{E}[\mathbf{R}_i]] \leq \frac{\mathbb{E}[\mathbf{R}_i]}{2\mathbb{E}[\mathbf{R}_i]} = \frac{1}{2}.$$

Not great...half the servers may be overloaded.

> n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

With a very simple twist Markov's Inequality can be made much more powerful.

CHEBYSHEV'S INEQUALITY

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

$\mathbf{X}^2$ is a nonnegative random variable. So can apply Markov's inequality:

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

$\mathbf{X}^2$ is a nonnegative random variable. So can apply Markov's inequality:

$$\Pr(\mathbf{X}^2 \geq t^2) \leq \frac{\mathbb{E}[\mathbf{X}^2]}{t^2}.$$

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

$\mathbf{X}^2$ is a nonnegative random variable. So can apply Markov's inequality:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2) \leq \frac{\mathbb{E}[\mathbf{X}^2]}{t^2}.$$

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

$\mathbf{X}^2$ is a nonnegative random variable. So can apply Markov's inequality:

Chebyshev's inequality:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2) \leq \frac{\mathbb{E}[\mathbf{X}^2]}{t^2}.$$

With a very simple twist Markov's Inequality can be made much more powerful.

For any random variable $\mathbf{X}$ and any value $t > 0$:

$$\Pr(|\mathbf{X}| \geq t) = \Pr(\mathbf{X}^2 \geq t^2).$$

$\mathbf{X}^2$ is a nonnegative random variable. So can apply Markov's inequality:

Chebyshev's inequality:

$$\Pr(|\mathbf{X} - \mathbb{E}[\mathbf{X}]| \geq t) \leq \frac{\mathrm{Var}[\mathbf{X}]}{t^2}.$$

(by plugging in the random variable $\mathbf{X} - \mathbb{E}[\mathbf{X}]$)

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\mathbf{R}_i = \sum_{j=1}^{n} \mathbf{R}_{i,j}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\text{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \text{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\mathrm{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathrm{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

$$\mathrm{Var}[\mathbf{R}_{i,j}] = \mathbb{E}\left[(\mathbf{R}_{i,j} - \mathbb{E}[\mathbf{R}_{i,j}])^2\right]$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\mathrm{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \mathrm{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

$$\begin{aligned}\mathrm{Var}[\mathbf{R}_{i,j}] &= \mathbb{E}\left[(\mathbf{R}_{i,j} - \mathbb{E}[\mathbf{R}_{i,j}])^2\right] \\ &= \Pr(\mathbf{R}_{i,j} = 1) \cdot (1 - \mathbb{E}[\mathbf{R}_{i,j}])^2 + \Pr(\mathbf{R}_{i,j} = 0) \cdot (0 - \mathbb{E}[\mathbf{R}_{i,j}])^2\end{aligned}$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\text{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \text{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

$$\begin{aligned}
\text{Var}[\mathbf{R}_{i,j}] &= \mathbb{E}\left[(\mathbf{R}_{i,j} - \mathbb{E}[\mathbf{R}_{i,j}])^2\right] \\
&= \Pr(\mathbf{R}_{i,j} = 1) \cdot (1 - \mathbb{E}[\mathbf{R}_{i,j}])^2 + \Pr(\mathbf{R}_{i,j} = 0) \cdot (0 - \mathbb{E}[\mathbf{R}_{i,j}])^2 \\
&= \frac{1}{k} \cdot \left(1 - \frac{1}{k}\right)^2 + \left(1 - \frac{1}{k}\right) \cdot \left(0 - \frac{1}{k}\right)^2
\end{aligned}$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\text{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \text{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

$$\begin{aligned}
\text{Var}[\mathbf{R}_{i,j}] &= \mathbb{E}\left[(\mathbf{R}_{i,j} - \mathbb{E}[\mathbf{R}_{i,j}])^2\right] \\
&= \Pr(\mathbf{R}_{i,j} = 1) \cdot (1 - \mathbb{E}[\mathbf{R}_{i,j}])^2 + \Pr(\mathbf{R}_{i,j} = 0) \cdot (0 - \mathbb{E}[\mathbf{R}_{i,j}])^2 \\
&= \frac{1}{k} \cdot \left(1 - \frac{1}{k}\right)^2 + \left(1 - \frac{1}{k}\right) \cdot \left(0 - \frac{1}{k}\right)^2 \\
&= \frac{1}{k} - \frac{1}{k^2} \le \frac{1}{k}
\end{aligned}$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

LOAD BALANCING VARIANCE

We can write the number of requests assigned to server i, $\mathbf{R}_i$ as:

$$\text{Var}[\mathbf{R}_i] = \sum_{j=1}^{n} \text{Var}[\mathbf{R}_{i,j}] \qquad \text{(linearity of variance)}$$

where $\mathbf{R}_{i,j}$ is 1 if request j is assigned to server i and 0 otherwise.

$$\begin{aligned}
\text{Var}[\mathbf{R}_{i,j}] &= \mathbb{E}\left[(\mathbf{R}_{i,j} - \mathbb{E}[\mathbf{R}_{i,j}])^2\right] \\
&= \Pr(\mathbf{R}_{i,j} = 1) \cdot (1 - \mathbb{E}[\mathbf{R}_{i,j}])^2 + \Pr(\mathbf{R}_{i,j} = 0) \cdot (0 - \mathbb{E}[\mathbf{R}_{i,j}])^2 \\
&= \frac{1}{k} \cdot \left(1 - \frac{1}{k}\right)^2 + \left(1 - \frac{1}{k}\right) \cdot \left(0 - \frac{1}{k}\right)^2 \\
&= \frac{1}{k} - \frac{1}{k^2} \le \frac{1}{k} \implies \text{Var}[\mathbf{R}_i] \le \frac{n}{k}.
\end{aligned}$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \leq \frac{n}{k}$.

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \leq \frac{n}{k}$.

Applying Chebyshev's:

$$\Pr\left(\mathbf{R}_i \geq \frac{2n}{k}\right) \leq \Pr\left(|\mathbf{R}_i - \mathbb{E}[\mathbf{R}_i]| \geq \frac{n}{k}\right)$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \le \frac{n}{k}$.

Applying Chebyshev's:

$$\Pr\left(\mathbf{R}_i \ge \frac{2n}{k}\right) \le \Pr\left(|\mathbf{R}_i - \mathbb{E}[\mathbf{R}_i]| \ge \frac{n}{k}\right) \le \frac{n/k}{n^2/k^2}$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \leq \frac{n}{k}$.

Applying Chebyshev's:

$$\Pr\left(\mathbf{R}_i \geq \frac{2n}{k}\right) \leq \Pr\left(|\mathbf{R}_i - \mathbb{E}[\mathbf{R}_i]| \geq \frac{n}{k}\right) \leq \frac{n/k}{n^2/k^2} = \frac{k}{n}.$$

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \leq \frac{n}{k}$.

Applying Chebyshev's:

$$\Pr\left(\mathbf{R}_i \geq \frac{2n}{k}\right) \leq \Pr\left(|\mathbf{R}_i - \mathbb{E}[\mathbf{R}_i]| \geq \frac{n}{k}\right) \leq \frac{n/k}{n^2/k^2} = \frac{k}{n}.$$

- Overload probability is extremely small when $k \ll n$!

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.

BOUNDING THE LOAD VIA CHEBYSHEVS

Letting $\mathbf{R}_i$ be the number of requests sent to server i, $\mathbb{E}[\mathbf{R}_i] = \frac{n}{k}$ and $\text{Var}[\mathbf{R}_i] \leq \frac{n}{k}$.

Applying Chebyshev's:

$$\Pr\left(\mathbf{R}_i \geq \frac{2n}{k}\right) \leq \Pr\left(|\mathbf{R}_i - \mathbb{E}[\mathbf{R}_i]| \geq \frac{n}{k}\right) \leq \frac{n/k}{n^2/k^2} = \frac{k}{n}.$$

- Overload probability is extremely small when $k \ll n$!
- Might seem counterintuitive – bound gets worse as k grows.
- When k is large, the number of requests each server sees in expectation is very small so the law of large numbers doesn't 'kick in'.

n: total number of requests, k: number of servers randomly assigned requests, $\mathbf{R}_i$: number of requests assigned to server i.