**BACKGROUND**

- **Streaming Algorithm:**
  - One pass over $N$ input elements
  - Maintain at most $o(N)$ elements in memory
  - Worst case/random stream order
  - Randomized/deterministic algorithm
  - Approximation ratio $\mathbb{E}[f(S)] \geq R \cdot f(OPT)$

- **Assumptions:**
  - Nonnegative $f(A) \geq 0, \forall A$
  - Monotone $f(B | A) \geq 0, \forall A, B$
  - $\gamma_k$-weakly submodular $\gamma_k \leq \min_{L \neq \emptyset} \frac{\sum_{j \in L} f(j | L)}{f(S) - f(L)}$

- **Example Function:**
  $$f_k(S) = \min(2 \cdot |S \cap U| + 1, 2 \cdot |S \cap V|)$$

- **U = \{u_1, \ldots, u_k\} V = \{v_1, \ldots, v_k\} D = \{w_1, \ldots, w_k\}$$

  - Worst case order begins with only elements from $U \cup D$
  - Sublinear streaming algorithms must drop many $u$ before any $v$ arrive
  - Approximation ratio is arbitrarily small for large $k$

- **Approximation Ratios:**
  - Let $\mathcal{E}$ be the event $f(S) < \tau$ (balanced if $f_2(\mathcal{E}) = 2 - \sqrt{2 - e^{-\gamma_k}}$)
  - $\mathbb{E}[f(S)] \geq (1 - Pr[\mathcal{E}]) \cdot \tau$
  - $\mathbb{E}[f(S)] \geq \frac{1}{2} \left( \gamma_k \cdot f_2(\mathcal{E}) - e^{-\gamma_k/2} \right)$, $f(OPT) - 2\tau$
  - Show one instance is guaranteed to be a good approximation

**PROOF TECHNIQUES**

- **Example Function:**
  $$f_k(S) = \min(2 \cdot |S \cap U| + 1, 2 \cdot |S \cap V|)$$

- **U = \{u_1, \ldots, u_k\} V = \{v_1, \ldots, v_k\} D = \{w_1, \ldots, w_k\}$$

  - Worst case order begins with only elements from $U \cup D$
  - Sublinear streaming algorithms must drop many $u$ before any $v$ arrive
  - Approximation ratio is arbitrarily small for large $k$

- **Approximation Ratios:**
  - Let $\mathcal{E}$ be the event $f(S) < \tau$ (balanced if $f_2(\mathcal{E}) = 2 - \sqrt{2 - e^{-\gamma_k}}$)
  - $\mathbb{E}[f(S)] \geq (1 - Pr[\mathcal{E}]) \cdot \tau$
  - $\mathbb{E}[f(S)] \geq \frac{1}{2} \left( \gamma_k \cdot f_2(\mathcal{E}) - e^{-\gamma_k/2} \right)$, $f(OPT) - 2\tau$
  - Show one instance is guaranteed to be a good approximation

**FUTURE WORK**

- Tighten approximation bounds
- Analyze additional classes of algorithms: randomized, $\gamma$ input
- Combinatorial interpretability for fairness, adversarial examples, ...

**SUMMARY**

Many discrete optimization applications have a very large ground set or an expensive function evaluation oracle. We design and analyze streaming algorithms for the general class of weakly submodular set functions:

- **Worst case stream order:** No randomized streaming algorithm using sublinear memory can maximize a 0.5-weakly submodular function with constant approximation ratio
- **Random stream order:** Greedy, deterministic streaming algorithm for weak submodular maximization with constant approximation ratio
- **Experimental Evaluation:** Nonlinear sparse regression and interpretability of black-box neural networks

**EXPERIMENTAL RESULTS**

Sparse logistic regression: Compute pairwise products of features as needed

- **Phishing Dataset, N=4.7k, 40 iterations**

Interpretability: Select image segments which maximize label’s likelihood

- **max \frac{softmax\_score(Image_2)}{\sum_{|S| \leq k}}**

Transfer Learning (InceptionV3 flower classification)

- **Comparison with LIME**

**REFERENCES**