ECE 606, Fall 2019, Assignment 8

Zhijie Wang, Student ID number: 20856733

zhijie.wang@uwaterloo.ca

November 5, 2019

1. (a) $\frac{|E|}{k}$.

   Let x_i represents the ith edge $\in E$, where $i = 1, 2, \ldots, |E|$. Then, we define the random variable X_i:

   $$X_i = \begin{cases} 1 & , x_i \text{ not satisfied} \\ 0 & , \text{otherwise} \end{cases}$$

   Hence, for each edge $x_i = (u, v)$, $Pr_\{X_i = 1\} = k/k^2 = \frac{1}{k}$. Therefore,

   $$E[X_i] = \sum_{i=1}^{|E|} \frac{1}{k} = \frac{|E|}{k}$$

   (b) Let k colours denote to $c_1, \ldots c_k$. Consider the colouring method that first colour all vertexes with c_1, then we randomly pick a vertex and recolour it with $c_2, \ldots c_k$ if the number of satisfied edges is increased, then go to another vertex till all the vertexes are recoloured. Now we prove this. Let the number of unsatisfied edges denote to S.

   Proof. Here we do a case analysis on k. 1) $k = 1$, and 2) $k \geq 2$. If $k = 1$, then $S = |E| = |E|/k$. Then, we consider about $k \geq 2$.

   Let v_i denote to the ith vertex, where $1 \leq i \leq |V|$. Now we prove the claim above through every sub-graph $G' =< V', E' >$ constructed by $v_1, \ldots v_i$.

   Base case: $i = 2$. In this sub-graph, there are only 1 edge, v_1 and v_2 cannot be the same colour by our colouring method, therefore, $S' = 0 < |E'|/k$

   Step: We assume that given any $i \in 1, \ldots, |v| - 1$, the sub-graph constructed by $v_1, \ldots v_i$ has a number of unsatisfied edges S' at most $|E'|/k$. Then we need to prove that after colouring $|V|$th vertex, $S = S' + \delta \leq |E|/k$, where $|E| = |E'| + d_{|V|}$, $d_{|V|}$ is the degree of $|V|$th vertex. Hence, we need to prove $\frac{|E| - d_{|V|}}{k} + \delta \leq \frac{|E|}{k} \Rightarrow \delta \leq \frac{d_{|V|}}{k}$, where δ is the increased unsatisfied edges.

   If $d_{|V|} < k$, according to our algorithm, after colouring the last vertex, the number of satisfied edges will increase by $d_{|V|}$, therefore, $\delta = 0$, then, $\frac{|E| - d_{|V|}}{k} + 0 < \frac{|E|}{k}$.

   If $d_{|V|} \geq k$, then, according to our algorithm, $\delta \leq \frac{d_{|V|}}{k}$. Here we prove this by contradiction. Let m_i denote to the number of vertexes incident on $|V|$th has a colour of c_i. If all $m_i > |E|/k$, then $\sum m_i > |E|$, which contradicts on $\sum m_i = |E|$. Therefore, there exists m_i which is not bigger than $|E|/k$. Then, $\delta \leq \frac{d_{|V|}}{k}$. Hence, $\frac{|E| - d_{|V|}}{k} + \delta \leq \frac{|E| - d_{|V|}}{k} + \frac{d_{|V|}}{k} = \frac{|E|}{k}$.

   Therefore, there exists a colouring that at most $|E|/k$ unsatisfied edges.

2. (a) *Proof.* For each key $k \in \mathcal{U}$, $h(k, i)$ is distinct. Therefore, if h_1 is a random function, which means $Pr\{h_1(k) = j\} = 1/m$, where $k \in \mathcal{U}$ and $j \in \{0, \ldots m - 1\}$, then, $Pr\{h_1(k) + (i - 1)h_2(k) = j'\} = Pr\{h_1(k) = j\} = 1/m$.

   Hence, $Pr\{h(k) = j''\} = Pr\{h_1(k) + (i - 1)h_2(k) = j'\} = Pr\{h_1(k) = j\} = 1/m$, where $j'' \in \{0, \ldots m - 1\}$. h is also a random function.

   (b) *Proof.* Let $h_2(k) = dc_1$ and $m = dc_2$ for constants c_1, c_2.

   Suppose we have examined $\frac{1}{d}$th of the table entries, which means $(i - 1) = \frac{m}{d} = c_2$. The probe that we check at this point is $[h_1(k) + (m/d)h_2(k)] \mod m = [h_1(k) + (m/d)(dc_1)] \mod m \equiv h_1(k)$. When $d = 1$, we will end up probing all slots in the table in the first m probes.

   Therefore, If $\gcd(h2(k), m) = 1$, we end up probing all slots in the table in the first m probes.

3. Here we can consider this problem as a version of set-covering problem. The universe $\mathcal{U}$ is the set E of edges of G. Each vertex v has a subset S_v which contains all the edges incident on v, therefore, these subsets can be considered as the family $\mathcal{F}$.

   Hence, since each edge $e = (u, v)$ would appear in both S_u and S_v, therefore, the union of all these subsets in $\mathcal{F}$ equals to $\mathcal{U}$.

   Now, because each time the algorithm always select the vertex with highest degree, which means the subset has a biggest size, this problem is the same as set-covering problem. As a result, the algorithm has a running-time of:

$$O(1 + \lg |E|) \leq O(\lg(n)^2) = O(2 \lg n)$$

   therefore, this algorithm achieves an approximation ratio of $O(\lg n)$, where n is the size of input.