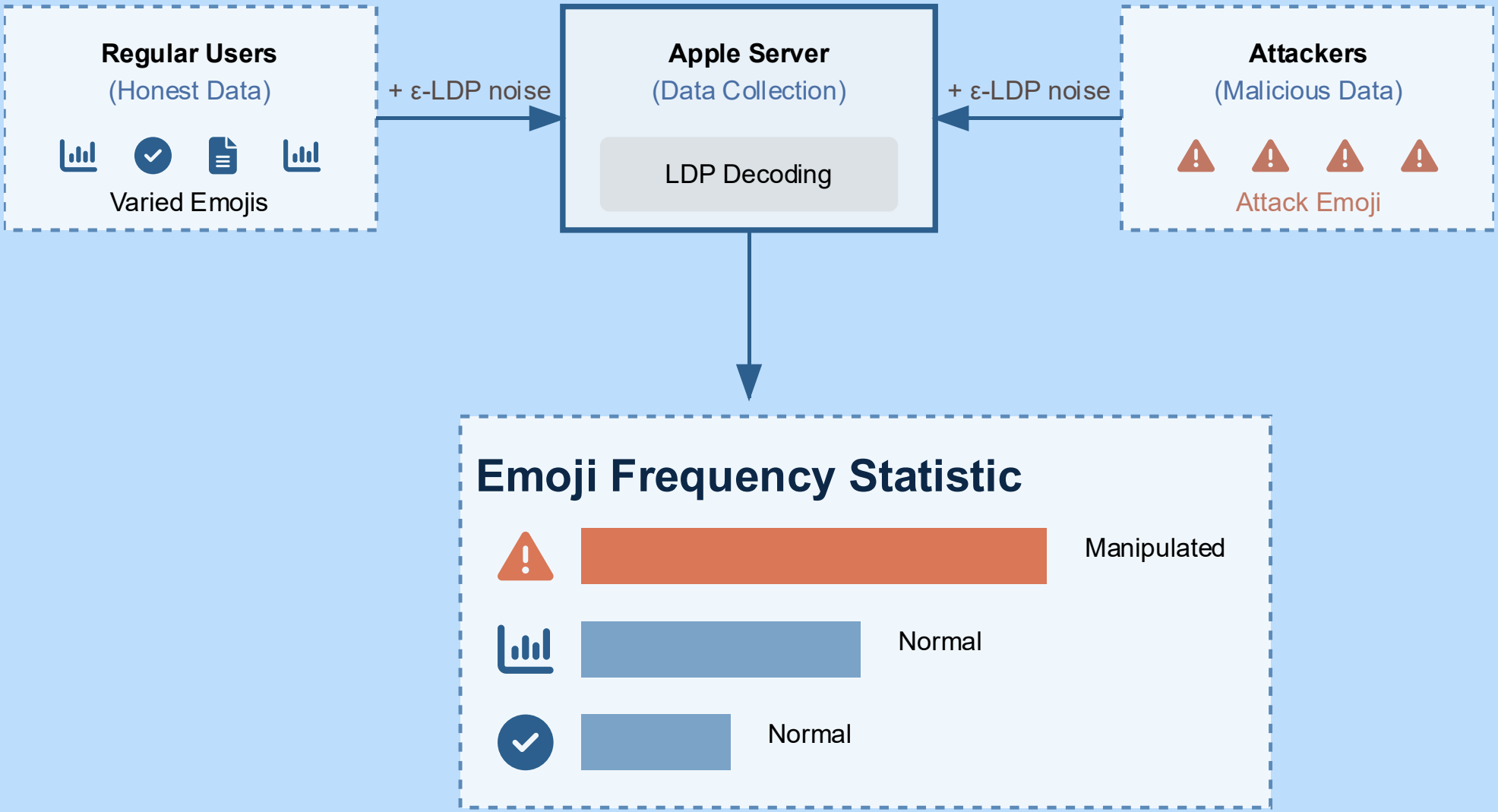
Maximum Gain Attacks on Frequency Estimators

LDP Protocol Attack Skewing the Frequently Used Emoji Statistic



General Formula for Maximum Gain

$$G = \frac{\sum_{\{i=n+1\}}^{\{n+m\}} \mathbb{E}\left[\mathbb{I}_{Support(y_i)}(t)\right]}{(n+m)(p^* - q^*)} - \frac{m \sum_{\{i=1\}}^{\{n\}} \mathbb{E}\left[\mathbb{I}_{Support(y_i)}(t)\right]}{n(n+m)(p^* - q^*)}$$

Attacking kRR

- The support set is:

$$\sum_{\{t \in T\}} \mathbb{I}_{Support(y_i)}(t) = 1$$

- Gain becomes:

$$G = \frac{m}{(n+m)(p^* - q^*)} - c$$

- Plugging in p^* and q^* gives:

$$G = \beta(1 - f_T) + \frac{\beta(d-r)}{e^{\epsilon} - 1}$$

Attacking OUE

- The support set is:

$$\sum_{\{t \in T\}} \mathbb{I}_{Support(y_i)}(t) = r$$

- Gain becomes:

$$G = \frac{rm}{(n+m)(p^* - q^*)} - c$$

- Plugging in p^* and q^* gives:

$$G = \beta(2r - f_T) + \frac{2\beta r}{e^{\epsilon} - 1}$$

Attacking OLH

- The support set is:

$$\sum_{\{t \in T\}} \mathbb{I}_{Support(y_i)}(t) = r$$

- Gain becomes:

$$G = \frac{rm}{(n+m)(p^* - q^*)} - c$$

- Plugging in p^* and q^* gives:

$$G = \beta(2r - f_T) + \frac{2\beta r}{e^{\epsilon} - 1}$$