Appendix

A Proof of Lemma 1

Now give out the proof of lemma 1, i.e.,

Lemma 1. *The following event, i.e.,*

$$\xi_1 = \{\forall(h,i) \in T_t, \forall t : |\hat{\mu}^0_{h,i}(t) - f(v_{h,i})| < B(T_{h,i}(t), \delta, t)\}$$

establishes with the probability at least $1 - \delta$

The User's rewards to video $v_{h,i}$ are a sequence of i.i.d(independently identically distribution) random variables and belong to [0,1], define event ξ_1^c is the opposite event of event ξ_1, then according to Hoeffding's inequality, we have, i.e.,

$$\begin{aligned}
P[\xi_1^c] &\leq \sum_{(h,i)\in T_t} \sum_{T_{h,i}(t)=1}^{\infty} 2exp\Big[\frac{-2T_{h,i}(t)^2}{T_{h,i}(t)\cdot 1^2} B(T_{h,i}(t), \delta, t)^2\Big] \\
&= \sum_{(h,i)\in T_t} \sum_{T_{h,i}(t)=1}^{\infty} 2exp\Big[log\frac{3\delta}{\pi^2 T_{h,i}(t)^2 |T_t|}\Big] \\
&= \sum_{(h,i)\in T_t} \sum_{T_{h,i}(t)=1}^{\infty} \frac{6\delta}{\pi^2 T_{h,i}(t)^2 |T_t|} \\
&= \sum_{(h,i)\in T_t} \frac{\delta}{|T_t|} \\
&= \delta.
\end{aligned}$$

so we have $P[\xi_1] > 1 - \delta$

B Proof of Lemma 2

The Lemma 2 is that, i.e.,

Lemma 2. *under event* ξ_1*, for* $\forall(h,i) \in T_t, \notin N_K(t), \forall(h_K, i_K) \in N_K(t), \forall t$*, there exists, i.e.,*

$$\hat{\mu}_{(h,i),(h_K,i_K)}(t) - f(v_{h_K,i_K}) < B(T_{(h,i),(h_K,i_K)}(t), \delta, t) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t)$$

$\bar{\lambda}_{(h,i),(h_K,i_K)}(t) = \frac{1}{T_{(h,i),(h_K,i_K)}(t)} \sum_{s\in\phi_{(h,i),(h_K,i_K)}(t)} \lambda_s$

The proof of it is as following,

$$
\begin{aligned}
&\hat{\mu}_{(h,i),(h_K,i_K)}(t) \\
&= \frac{1}{T_{(h,i),(h_K,i_K)}(t)} \sum_{s\in\phi_{(h,i),(h_K,i_K)}(t)} \left(r_s - \left[\hat{\mu}_{h,i}(s) - \hat{\mu}_{h_K,i_K}(s) + B(T_{h,i}(s),\delta,s) + B(T_{h_K,i_K}(s),\delta,s) + \lambda_s\right]^+\right) \\
&\leq \frac{1}{T_{(h,i),(h_K,i_K)}(t)} \sum_{s\in\phi_{(h,i),(h_K,i_K)}(t)} \left(r_s - \left[\hat{\mu}_{h,i}(s) - \hat{\mu}_{h_K,i_K}(s) + B(T_{h,i}(s),\delta,s) + B(T_{h_K,i_K}(s),\delta,s) + \lambda_s\right]\right) \\
&\overset{(i)}{<} \frac{1}{T_{(h,i),(h_K,i_K)}(t)} \sum_{s\in\phi_{(h,i),(h_K,i_K)}(t)} \left(r_s - (f(v_{h,i}) - f(v_{h_K,i_K}) + \lambda_s)\right) \\
&= \hat{\mu}^0_{(h,i),(h_K,i_K)}(t) - f(v_{h,i}) + f(v_{h_K,i_K}) - \frac{1}{T_{(h,i),(h_K,i_K)}(t)} \sum_{s\in\phi_{(h,i),(h_K,i_K)}(t)} \lambda_s \\
&= \hat{\mu}^0_{(h,i),(h_K,i_K)}(t) - f(v_{h,i}) + f(v_{h_K,i_K}) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t) \\
&\overset{(ii)}{<} f(v_{h,i}) + B(T_{(h,i),(h_K.i_K)}(t),\delta,t) - f(v_{h,i}) + f(v_{h_K,i_K}) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t) \\
&= B(T_{(h,i),(h_K.i_K)}(t),\delta,t) + f(v_{h_K,i_K}) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t) \rightarrow \\
&\hat{\mu}_{(h,i),(h_K,i_K)}(t) - f(v_{h_K,i_K}) < B(T_{(h,i),(h_K.i_K)}(t),\delta,t) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t)
\end{aligned}
$$

Lemma 2 has been proofed. (i) is because the event ξ_1, and (ii) is because $\{r_s\}_{s\in\phi_{(h,i),(h_K,i_K)}(t)}$ is also a sequence of i.i.d variables and we can also use ξ_1 to deal with it.

C Proof of Lemma 3

The Lemma 3 is that, i.e.,

Lemma 3. *Under Lemma 2, for* $\forall (h,i) \in T_t, \notin N_K(t), \forall (h_K,i_K) \in N_K(t), \forall t$*, there exists,i.e.,*

$$
\hat{\mu}_{h,i}(t) < \frac{1}{T_{h,i}(t)} \sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) + B(\frac{T_{h,i}(t)}{|N_K(t)|},\delta,t) - \bar{\lambda}_{h,i}(t)
$$

$\bar{\lambda}_{h,i}(t) = \frac{1}{T_{h,i}(t)} \sum_{s\in\phi_{h,i}(t)} \lambda_s$

The proof is as following,

$$
\begin{aligned}
&\hat{\mu}_{h,i}(t) \\
&= \frac{1}{T_{h,i}(t)} \sum_{(h_K,i_K)\in N_K(t)} \hat{\mu}_{(h,i),(h_K,i_K)}(t) \cdot T_{(h,i),(h_K,i_K)}(t) \\
&\overset{(i)}{<} \frac{1}{T_{h,i}(t)} \sum_{(h_K,i_K)\in N_K(t)} \left(f(v_{h_K,i_K}) + B(T_{(h,i),(h_K,i_K)}(t),\delta,t) - \bar{\lambda}_{(h,i),(h_K,i_K)}(t)\right) T_{(h,i),(h_K,i_K)}(t) \\
&= \frac{1}{T_{h,i}(t)} \sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) + \frac{1}{T_{h,i}(t)} \sum_{(h_K,i_K)\in N_K(t)} T_{(h,i),(h_K,i_K)}(t) B(T_{(h,i),(h_K,i_K)}(t),\delta,t) - \bar{\lambda}_{h,i}(t) \\
&= \frac{1}{T_{h,i}(t)} \sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) + \frac{1}{T_{h,i}(t)} \sum_{(h_K,i_K)\in N_K(t)} \sqrt{\frac{T_{(h,i),(h_K,i_K)}(t)}{2} ln \frac{\pi^2 T_{(h,i),(h_K,i_K)}(t)^2 |T_t|}{3\delta}} - \bar{\lambda}_{h,i}(t)
\end{aligned}
\tag{1}
$$

(i) is under Lemma 2. we define a function $g(N) = \sqrt{\frac{N}{2} ln \frac{\pi^2 N^2 |T_t|}{3\delta}} (N \geq 1)$, because there exists $g(N)'' < 0$, so g(N) is a Concave Function, then we have, i.e.,

$$\begin{aligned} &\sum_{(h_K, i_K) \in N_K(t)} g(T_{(h,i),(h_K,i_K)}(t)) \\ &< |N_K(t)| g\big(\frac{1}{|N_K(t)|} \sum_{(h_K,i_K) \in N_K(t)} T_{(h,i),(h_K,i_K)}(t)\big) \\ &= |N_K(t)| g(\frac{1}{|N_K(t)|} T_{h,i}(t)) \end{aligned}$$

Based on the above, we continue the proof with (1),

$$\begin{aligned} (1) &< \frac{1}{T_{h,i}(t)} \sum_{s \in \phi_{h,i}(t)} f(v_{h_{Ks}, i_{Ks}}) + \frac{1}{T_{h,i}(t)} \sqrt{\frac{|N_K(t)| \cdot T_{h,i}(t)}{2} ln \frac{\pi^2 T_{h,i}(t)^2 |T_t|}{3\delta \cdot |N_K(t)|^2}} - \bar{\lambda}_{h,i}(t) \\ &= \frac{1}{T_{h,i}(t)} \sum_{s \in \phi_{h,i}(t)} f(v_{h_{Ks}, i_{Ks}}) + B(\frac{T_{h,i}(t)}{|N_K(t)|}, \delta, t) - \bar{\lambda}_{h,i}(t) \end{aligned}$$

The Lemma 3 has been proofed.

D Proof of Lemma 4

The Lemma 4 is about the height of cover-tree T_t, i.e.,

Lemma 4.

$$H(t) \leq H(t)_{max} < log_m \big[\frac{\nu_1^2 (1-\rho^2) t}{c^2} + 1\big] + 1$$

$m = \rho^{-2}$, $c = 2\sqrt{1/(1-\rho)}$

according to the HCT algorithm, a leaf node (h,i) is expanded when $\nu_1 \rho^h \geq c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}}$, so we have, i.e.,

$$T_{h,i}(t) \geq \frac{c^2 ln(1/\tilde{\delta}(t^+))}{\nu_1^2} \rho^{-2h} \geq \frac{c^2}{\nu_1^2} \rho^{-2h}$$

absolutely, when the tree is a linear tree, i.e. at each depth, only one node is been expanded, the tree is the deepest, so there exists,

$$\begin{aligned} T &\geq \sum_{h=1}^{H(T)-1} \frac{c^2}{\nu_1^2} \rho^{-2h} = \frac{c^2}{\nu_1^2} \sum_{h=1}^{H(T)-1} \rho^{-2h} \\ &= \frac{c^2}{\nu_1^2} \rho^{-2} \frac{1-\rho^{-2(H(T)-1)}}{1-\rho^{-2}} \\ &= \frac{c^2}{\nu_1^2} \frac{\rho^{-2(H(T)-1)} - 1}{1-\rho^2} \rightarrow \\ &\rho^{-2(H(T)-1)} - 1 \leq \frac{\nu_1^2 (1-\rho^2) T}{c^2} \rightarrow \\ &H(T) \leq log_{\rho^{-2}} \big[\frac{\nu_1^2 (1-\rho^2) T}{c^2} + 1\big] + 1 \end{aligned}$$

Lemma 4 has been proofed.

E Proof of Lemma 5

The Lemma 5 is about the node number of cover-tree T_t, i.e.,

Lemma 5.

$$|T_t| \leq |T_t|_{max} < 4\big(t\nu_1^2(2-\rho^2)/(2c^2)+1\big)^E - 1$$

$E = log_{2\rho^{-2}}2$.

As in Lemma 4, a node will be expanded until $T_{h,i}(t) \geq \frac{c^2}{\nu_1^2}\rho^{-2h}$, so the height is bigger, the threshold is bigger, absolutely, when the cover-tree is a Complete Binary Tree, is has the max node number, then we have, i.e.,

$$\begin{aligned}
T &\geq \sum_{h=1}^{H(T)-1} \frac{c^2}{\nu_1^2}\rho^{-2h}\cdot 2^h = \sum_{h=1}^{H(T)-1} \frac{c^2}{\nu_1^2}(2\rho^{-2})^h \\
&= 2\frac{c^2}{\nu_1^2}\frac{(2\rho^{-2})^{H(T)-1}-1}{2-\rho^2} \rightarrow \\
&(2\rho^{-2})^{H(T)-1} - 1 \leq \frac{T\nu_1^2(2-\rho^2)}{2c^2} \rightarrow \\
&H(T) \leq log_{2\rho_{-2}}\big[\frac{T\nu_1^2(2-\rho^2)}{2c^2}+1\big]+1
\end{aligned}$$

Then we use $2^{H(T)+1}-1$ to calculate the node number and get Lemma 5.

F Proof of Theorem 1

The Theorem 1 implies the higher-bound of the cost, i.e.,

Theorem 1.

$$\begin{aligned}
C(T) &\leq 2(f_{max}-f_{min}+4B(1,\delta,T))|T_T|\frac{\Big(\sqrt{\frac{|N_K(T)|}{2}ln\frac{\pi^2T^2|T_T|}{3\delta\cdot|N_K(T)|^2}}+2c\sqrt{ln(1/\tilde{\delta}(T^+))}\Big)^2}{\alpha_T^2} \\
&= O(\frac{1}{\alpha_T^2}(lnT)^3T^E)
\end{aligned}$$

establishes with the probability at least $1-\delta$*,* $E = log_{2\rho^{-2}}2$*,* $T^+ = 2^{\lfloor lnT\rfloor+1}$*,* $\tilde{\delta}(t) = min\{c_1\delta_u/t, 1\}(c_1 = \sqrt[8]{\rho/(3\nu_1)})$ *and* $\alpha_T > \min_{(h_a,i_a),(h_b,i_b)\in T_T}\{|f(v_{h_a,i_a})-f(v_{h_b,i_b})|\}$.

First, we make an assumption that at round t, the user chooses a node (h,i) excluding the target video v_K along a path P_t, in the path, we have, i.e.,

$$B_{h',i'}(t) \leq U_{h,i}(t)(h'<h,(h',i')\in P_t). \tag{2}$$

Because root node contains video v_K, so along the path, there must be a node $(h_K,i_K)(h_K<h)$ containing video v_K. At the same time, because when the user chooses a node containing video v_K, the attacker won't attack, so we still can use the property of the typical HCT algorithm, i.e. event ξ_t at Lemma 3 in [azar, M.G., Lazaric, A. Brunskill, E.. (2014). Online Stochastic Optimization under Correlated Bandit Feedback.[C] Proceedings of the

31st In-ternational Conference on Machine Learning, in PMLR 32(2):1557-1565] to analyze it, then under ξ_t, we have, i.e.,

$$
\begin{aligned}
U_{h_K,i_K}(t) &= \widehat{\mu}_{h_K,i_K}(t) + \nu_1\rho^{h_K} + \sqrt{\frac{c^2\log(1/\tilde{\delta}(t^+))}{T_{h_K,i_K}(t)}} \\
&\overset{(i)}{\geq} f(x_{h_K,i_K}) + \nu_1\rho^{h_K} \\
&\geq f(v_K).
\end{aligned}
$$

(i) is because under the event ξ_t ($P[\xi_t] \geq 1-\delta$). For the leaf node (h_n, i_n) containing video v_K, obviously, we have, i.e.,

$$
B_{h_n,i_n}(t) = U_{h_n,i_n}(t) \geq f(v_K),
$$

also according to the definition of B-value in ***HCT***, we have

$$
B_{h_K,i_K}(t) = \min\big[U_{h_K,i_K}(t), \max_{j\in\{2i_K-1,2i_K\}} B_{h_K+1,j}(t)\big], \tag{3}
$$

established, and between nodes $(h_K+1, 2i_K-1)$ and $(h_K+1, 2i_K)$, there must have a node containing video v_K, also node (h_K, i_K) must be the ancestor of node (h_n, i_n), now by propagating the bound backward from (h_n, i_n) to (h_K, i_K) through the (3) we can show that $B_{h_K,i_K}(t)$ is still a valid upper bound of $f(v_K)$.

Then from Inq.(2), we have, i.e.,

$$
\begin{aligned}
& U_{h,i}(t) \geq B_{h_K,i_K}(t) > f(v_K) \rightarrow \\
& \hat{\mu}_{h,i}(t) + \nu_1\rho^h + c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}} \geq f(v_K) \overset{(i)}{\rightarrow} \\
& f(v_K) \leq \frac{1}{T_{h,i}(t)}\sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) + B(\frac{T_{h,i}(t)}{|N_K(t)|},\delta,t) - \bar{\lambda}_{h,i}(t) + \nu_1\rho^h + c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}} \\
& \qquad < \frac{1}{T_{h,i}(t)}\sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) + B(\frac{T_{h,i}(t)}{|N_K(t)|},\delta,t) - \bar{\lambda}_{h,i}(t) + 2c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}} \rightarrow \\
& f(v_K) + \bar{\lambda}_{h,i}(t) - \frac{1}{T_{h,i}(t)}\sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}}) < B(\frac{T_{h,i}(t)}{|N_K(t)|},\delta,t) + 2c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}} \rightarrow \\
& \quad \alpha_t < B(\frac{T_{h,i}(t)}{|N_K(t)|},\delta,t) + 2c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}}\big(\alpha_t = f(v_K) + \bar{\lambda}_{h,i}(t) - \frac{1}{T_{h,i}(t)}\sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks},i_{Ks}})\big) \\
& \qquad = \sqrt{\frac{|N_K(t)|}{2T_{h,i}(t)} ln\frac{\pi^2 T_{h,i}(t)^2|T_t|}{3\delta\cdot|N_K(t)|^2}} + 2c\sqrt{\frac{ln(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}}
\end{aligned}
$$

We regard $T_{h,i}(t)$ as the unknown, then deal with the inequality and get, i.e.,

$$
T_{h,i}(t) < \frac{\sqrt{\frac{|N_K(t)|}{2} ln\frac{\pi^2 T_{h,i}(t)^2|T_t|}{3\delta\cdot|N_K(t)|^2}} + 2c\sqrt{ln(1/\tilde{\delta}(t^+))}}{\alpha_t^2}
$$

We assume that the node excluding video v_K have been selected for $A(t)$ times until round t, the we have, i.e.,

$$\begin{aligned}
A(T) &= \sum_{(h,i)\in T_T, v_K \notin \mathcal{P}_{h,i}} T_{h,i}(T) \\
&< \sum_{(h,i)\in T_T, v_K \notin \mathcal{P}_{h,i}} \frac{\sqrt{\frac{|N_K(T)|}{2} ln \frac{\pi^2 T_{h,i}(T)^2 |T_T|}{3\delta \cdot |N_K(T)|^2}} + 2c\sqrt{ln(1/\tilde{\delta}(T^+))}}{\alpha_T^2} \\
&< |T_T| \frac{\left(\sqrt{\frac{|N_K(T)|}{2} ln \frac{\pi^2 T^2 |T_T|}{3\delta \cdot |N_K(T)|^2}} + 2c\sqrt{ln(1/\tilde{\delta}(T^+))}\right)^2}{\alpha_T^2}
\end{aligned} \tag{4}$$

Because $v_K \in (h_{Ks}, i_{Ks})$, so approximately, we think that $f(v_K) \approx \frac{1}{T_{h,i}(t)} \sum_{s\in\phi_{h,i}(t)} f(v_{h_{Ks}, i_{Ks}})$, then $\alpha_t \approx \bar{\lambda}_{h,i}(t) > \min_{(h_a,i_a),(h_b,i_b)\in T_t} \{|f(v_{h_a,i_a}) - f(v_{h_b,i_b})|\}$. And at the same time, we define the function $h(N) = B(N, \delta, t)(N \geq 1)$, absolutely, function $h(N)$ is a decreasing function, so $h(1) \geq h(N)$. Then for $\forall (h_a, i_a), (h_b, i_b) \in T_t$, there exists, i.e.,

$$\begin{aligned}
&\hat{\mu}^0_{h_a,i_a}(t) - \hat{\mu}^0_{h_b,i_b}(t) + B(T_{h_a,i_a}(t), \delta, t) + B(T_{h_b,i_b}(t), \delta, t) \\
&< \left(\hat{\mu}^0_{h_a,i_a}(t) - B(T_{h_a,i_a}(t), \delta, t)\right) - \left(\hat{\mu}^0_{h_b,i_b}(t) + B(T_{h_b,i_b}(t), \delta, t)\right) + 2B(T_{h_a,i_a}(t), \delta, t) + 2B(T_{h_b,i_b}(t), \delta, t) \\
&\overset{(i)}{<} f(v_{h_a,i_a}) - f(v_{h_b,i_b}) + 2B(T_{h_a,i_a}(t), \delta, t) + 2B(T_{h_b,i_b}(t), \delta, t) \\
&< f(v_{h_a,i_a}) - f(v_{h_b,i_b}) + 4B(1, \delta, t) \\
&< f_{max} - f_{min} + 4B(1, \delta, t)
\end{aligned}$$

(i) is under event ξ_1.

Then we give out a higher-bound of η_t, i.e.,

$$\eta_t \leq -I\{(h_t, i_t) \notin N_K(t)\}\big(2(f_{max} - f_{min} + 4B(1, \delta, t))\big) \tag{5}$$

Finally, conbine Inq.(4), Inq.(5), Lemma 4 and Lemma 5, we get Theorem 1.

G Proof of Theorem 2

The Theorem gives out a lower-bound of HCT algorithm's regret under the proposed attack, i.e.,

Theorem 2.

$$\begin{aligned}
R(T) &> \Omega\big[(f_{max} - f(v_K))A - (ln(A/\delta_u))^{1/(d+2)} A^{(d+1)/(d+2)} - \sqrt{2Aln(A/\delta)}\big] \\
&\quad + \Omega\big[(f_{max} - f(v_K))B - (ln(B/\delta_u))^{1/(d+2)} B^{(d+1)/(d+2)} - \sqrt{2Bln(B/\delta)}\big] \\
&= \Omega((f_{max} - f(v_K))T)
\end{aligned}$$

with probability at least $(1 - \delta_u)(1 - \delta)$.

According to the definition of regret, we have, i.e.,

$$\begin{aligned}
R(T) &= Tf_{max} - \sum_{t=1}^{T} r_t \\
&= \sum_{t=1}^{T}(f_{max} - r_t) \\
&= \sum_{s\in\mathcal{A}^c}(f_{max} - r_s) + \sum_{s\in\mathcal{A}}(f_{max} - r_s) \\
&= \widehat{R}(T) + \tilde{R}(T)
\end{aligned}$$

We begin to deal with $\widehat{R}(T)$, in $\mathcal{A}^c$, the attacker chooses not to attack, which means that the user chooses a node (h, i) containing video v_K. Firstly, we make some transformations to $\widehat{R}(T)$, i.e.,

$$\begin{aligned}
\widehat{R}(T) &= \sum_{s\in\mathcal{A}^c}(f_{max} - r_s) \\
&= \sum_{s\in\mathcal{A}^c}(f_{max} - f(v_{h_s,i_s}) + f(v_{h_s,i_s}) - r_s) \\
&= \sum_{s\in\mathcal{A}^c}[f_{max} - f(v_K) + f(v_K) - f(v_{h_s,i_s}) + f(v_{h_s,i_s}) - r_s] \\
&= \sum_{s\in\mathcal{A}^c}(f_{max} - f(v_K)) - \sum_{s\in\mathcal{A}^c}(f(v_{h_s,i_s}) - f(v_K)) - \sum_{s\in\mathcal{A}^c}(r_s - f(v_{h_s,i_s})) \\
&= \sum_{s\in\mathcal{A}^c}(f_{max} - f(v_K)) - (a) - (b).
\end{aligned} \tag{6}$$

For (b), because $\{f(v_{h_s,i_s}) - r_s\}_{s\in\mathcal{A}^c}$ is a bounded martingale difference sequence and we have $|f(v_{h_s,i_s}) - r_s| \le 1$, so according to Azuma's inequality, we leads to, i.e.,

$$(b) = \sum_{s\in\mathcal{A}^c}(r_s - f(x_{h_s,i_s})) \le \sqrt{2Blog(B/\delta)}. \tag{7}$$

with probability at least $1 - \delta/B$.

then for (a), we have, i.e.,

$$\begin{aligned}
(a) &= \sum_{s\in\mathcal{A}^c}(f(v_{h_s,i_s}) - f(v_K)) \\
&= \sum_{(h,i)\in T_t}\sum_{s\in\mathcal{A}^c}(f(v_{h,i}) - f(v_K))I_{(h_s,i_s)=(h,i)} \\
&\le \sum_{(h,i)\in T_t}\sum_{s\in\mathcal{A}^c}(\nu_1\rho^h)I_{(h_s,i_s)=(h,i)} \\
&\overset{(i)}{\le} \sum_{(h,i)\in T_t}\sum_{s\in\mathcal{A}^c} c\sqrt{\frac{log(1/\tilde{\delta}(s^+))}{T_{h,i}(s)}}I_{(h_s,i_s)=(h,i)} \\
&\overset{(ii)}{\le} \left(\frac{2^{2(d+3)}\nu_1^{2(d+1)}C\nu_2^{-d}\rho^d}{(1-\rho)^{d/2+3}}\right)^{\frac{1}{d+2}}\left(\log\left(\frac{2B}{\delta_u}\sqrt[8]{\frac{3\nu_1}{\rho}}\right)\right)^{\frac{1}{d+2}}B^{\frac{d+1}{d+2}}.
\end{aligned} \tag{8}$$

(i) is because the property of the $OptTraverse$ function in HCT that the loop in the function will ends with a node meeting the condition, i.e.,

$$\nu_1 \rho^h < c\sqrt{\frac{\log(1/\tilde{\delta}(t^+))}{T_{h,i}(t)}}. \tag{9}$$

(ii) is from the Theorem 1 in [azar, M.G., Lazaric, A. Brunskill, E.. (2014). Online Stochastic Optimization under Correlated Bandit Feedback.[C] Proceedings of the 31st In-ternational Conference on Machine Learning, in PMLR 32(2):1557-1565], and with a probability at least $1 - \delta_u$.

Then combine (6),(7) and (8), we have, i.e.,

$$\begin{aligned}\widehat{R}(T) \geq B(f_{max} - f(v_K)) - \sqrt{2Blog(B/\delta)} - \\ \left(\frac{2^{2(d+3)}\nu_1^{2(d+1)}C\nu_2^{-d}\rho^d}{(1-\rho)^{d/2+3}}\right)^{\frac{1}{d+2}}\left(\log\left(\frac{2B}{\delta_u}\sqrt[8]{\frac{3\nu_1}{\rho}}\right)\right)^{\frac{1}{d+2}} B^{\frac{d+1}{d+2}}.\end{aligned} \tag{10}$$

Then we analyze $\tilde{R}(T)$, i.e.,

$$\begin{aligned}\tilde{R}(T) &= \sum_{s\in\mathcal{A}}(f_{max} - r_s) \\ &\geq \sum_{s\in\mathcal{A}}\left(f_{max} - (r_s^0 - \hat{\mu}^0_{h_s,i_s}(s) + \hat{\mu}^0_{h_{Ks},i_{Ks}}(s) - B(T_{h_s,i_s}(s),\delta,s) - B(T_{h_{Ks},i_{Ks}}(s),\delta,s) - \lambda_s)\right) \\ &\overset{(i)}{\geq} \sum_{s\in\mathcal{A}}\left(f_{max} - (r_s^0 - f(v_{h_s,i_s}) + f(v_{h_{Ks},i_{Ks}}) - \lambda_s)\right) \\ &\geq \sum_{s\in\mathcal{A}}\left[f_{max} - f(v_{h_{Ks},i_{Ks}}) - (r_s^0 - f(v_{h_s,i_s})) + \lambda_s\right] \\ &= \sum_{s\in\mathcal{A}}\left[f_{max} - f(v_K) + [\lambda_s - (f(v_{h_{Ks},i_{Ks}}) - f(v_K))] - (r_s^0 - f(v_{h_s,i_s})\right] \\ &= \sum_{s\in\mathcal{A}}\left[f_{max} - f(v_K)\right] + (c) - (d)\end{aligned} \tag{11}$$

For (d), similarly to (b), according to Azuma's inequality, we leads to, i.e.,

$$(d) = \sum_{s\in\mathcal{A}}(r_s^0 - f(v_{h_s,i_s})) \leq \sqrt{2Alog(A/\delta)} \tag{12}$$

with probability at least $1 - \delta/A$.

for (c), according to the definition of λ_s and the fact that $v_{h_{Ks},i_{Ks}}, v_K \in \mathcal{P}_{h_{Ks},i_{Ks}}$, so there exists, i.e.,

$$(c) = \sum_{s\in\mathcal{A}}\lambda_s - (f(v_{h_{Ks},i_{Ks}}) > 0$$

Then combine inq.(11), inq.(12), we have, i.e.,

$$\tilde{R}(T) > A(f_{max} - f(v_K)) - \sqrt{2Alog(A/\delta)} \tag{13}$$

Finally, we combine Inq.(10) and Ieq.(13) to get Theorem 2.

H Additional Experiments

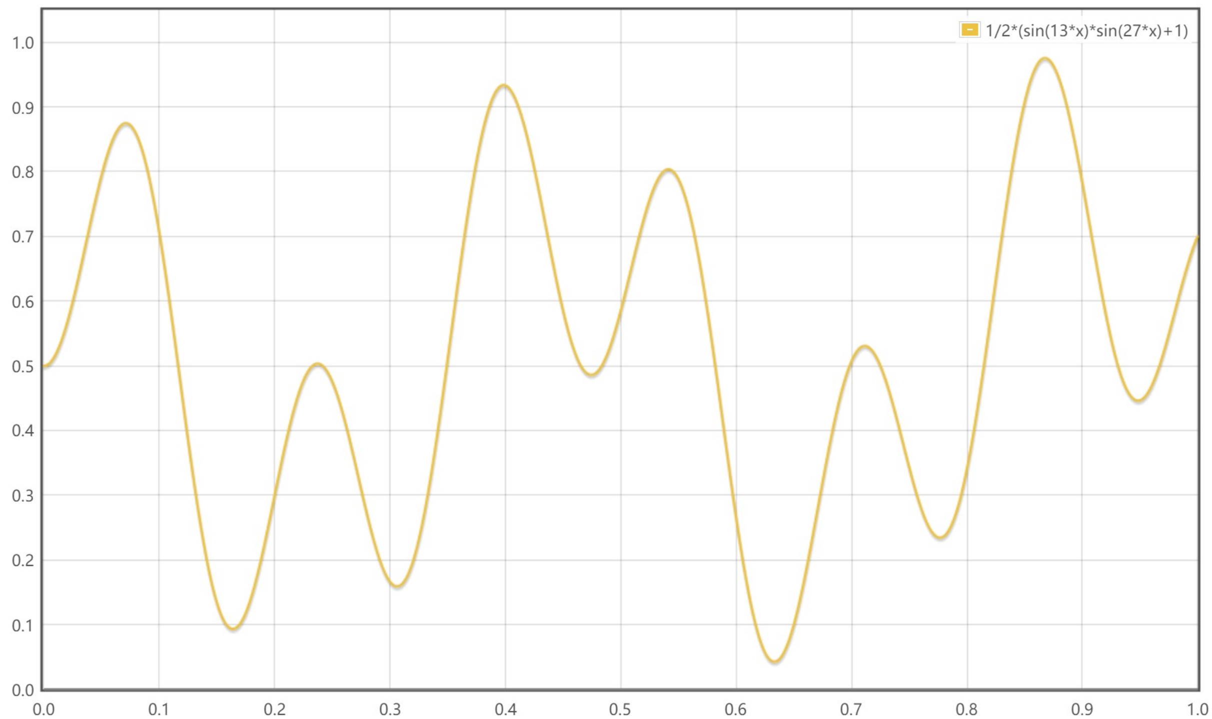
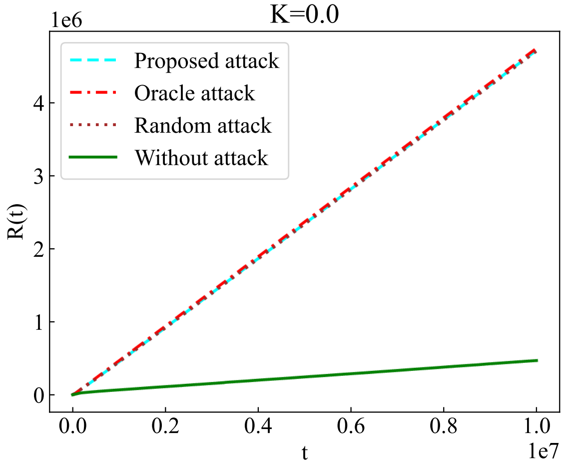
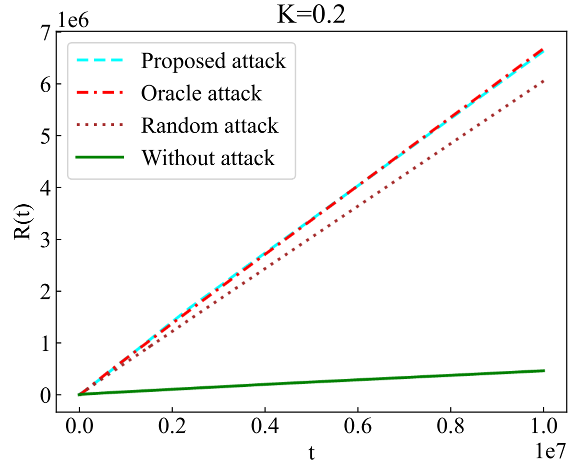


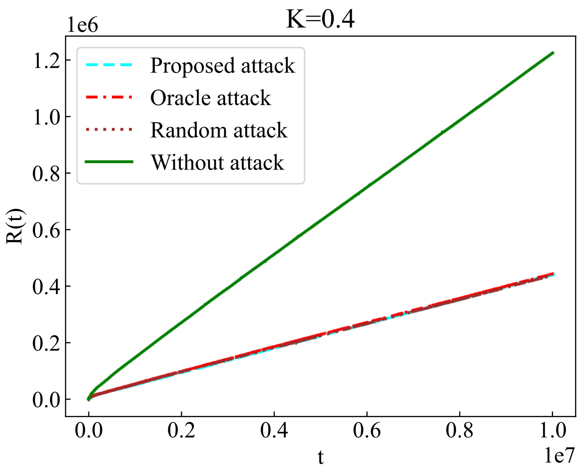
Figure 1: $f(x) = 1/2(sin(13x)sin(27x) + 1)$



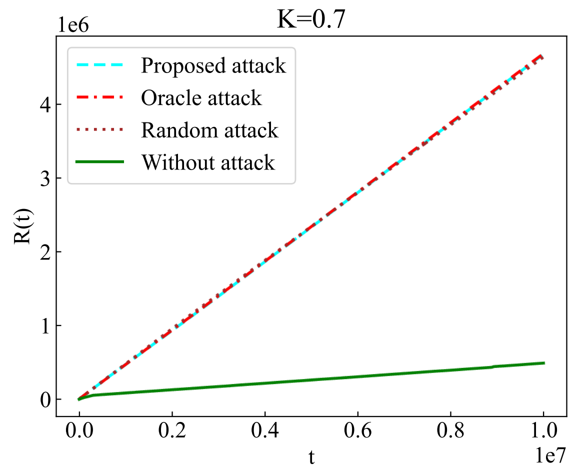
(a) $|v_K| = 0.0$



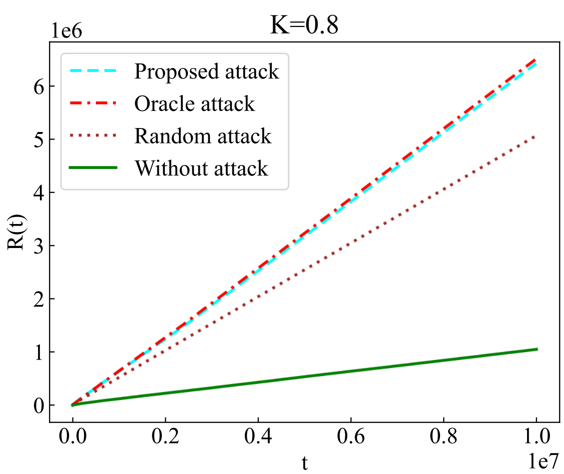
(b) $|v_K| = 0.2$



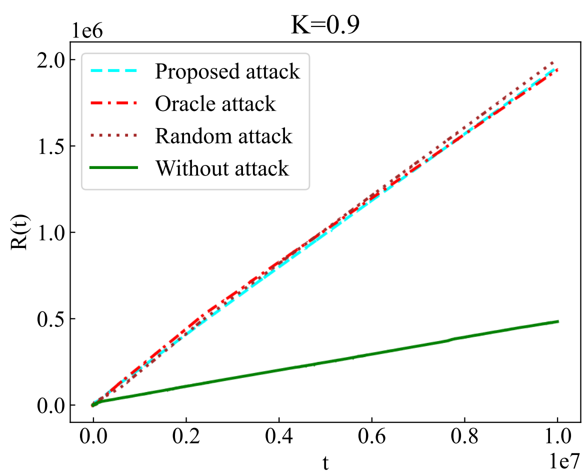
(c) $|v_K| = 0.4$



(d) $|v_K| = 0.7$

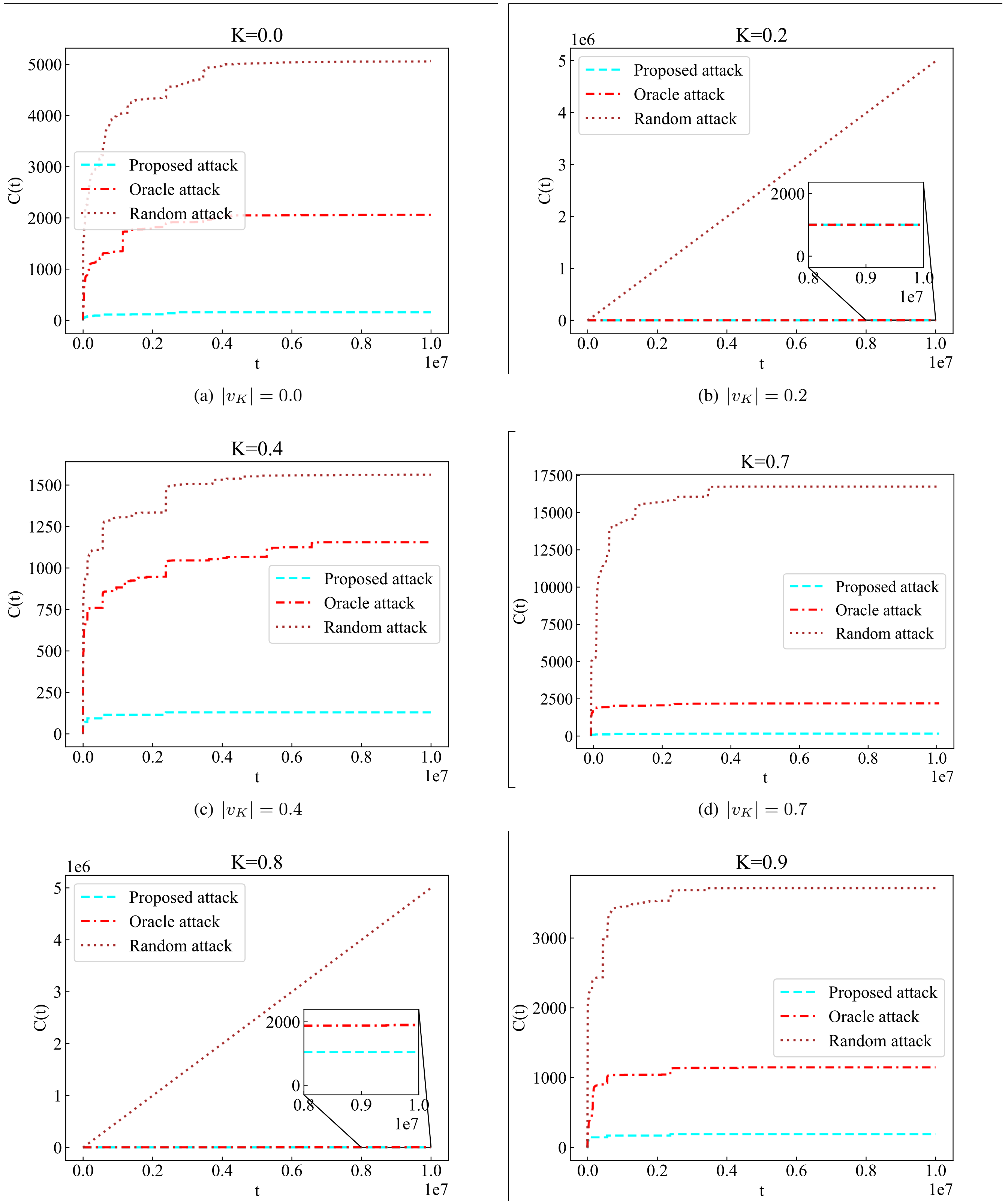


(e) $|v_K| = 0.8$



(f) $|v_K| = 0.9$

Figure 2: The graphs of $R(t) - t$



(a) $|v_K| = 0.0$

(b) $|v_K| = 0.2$

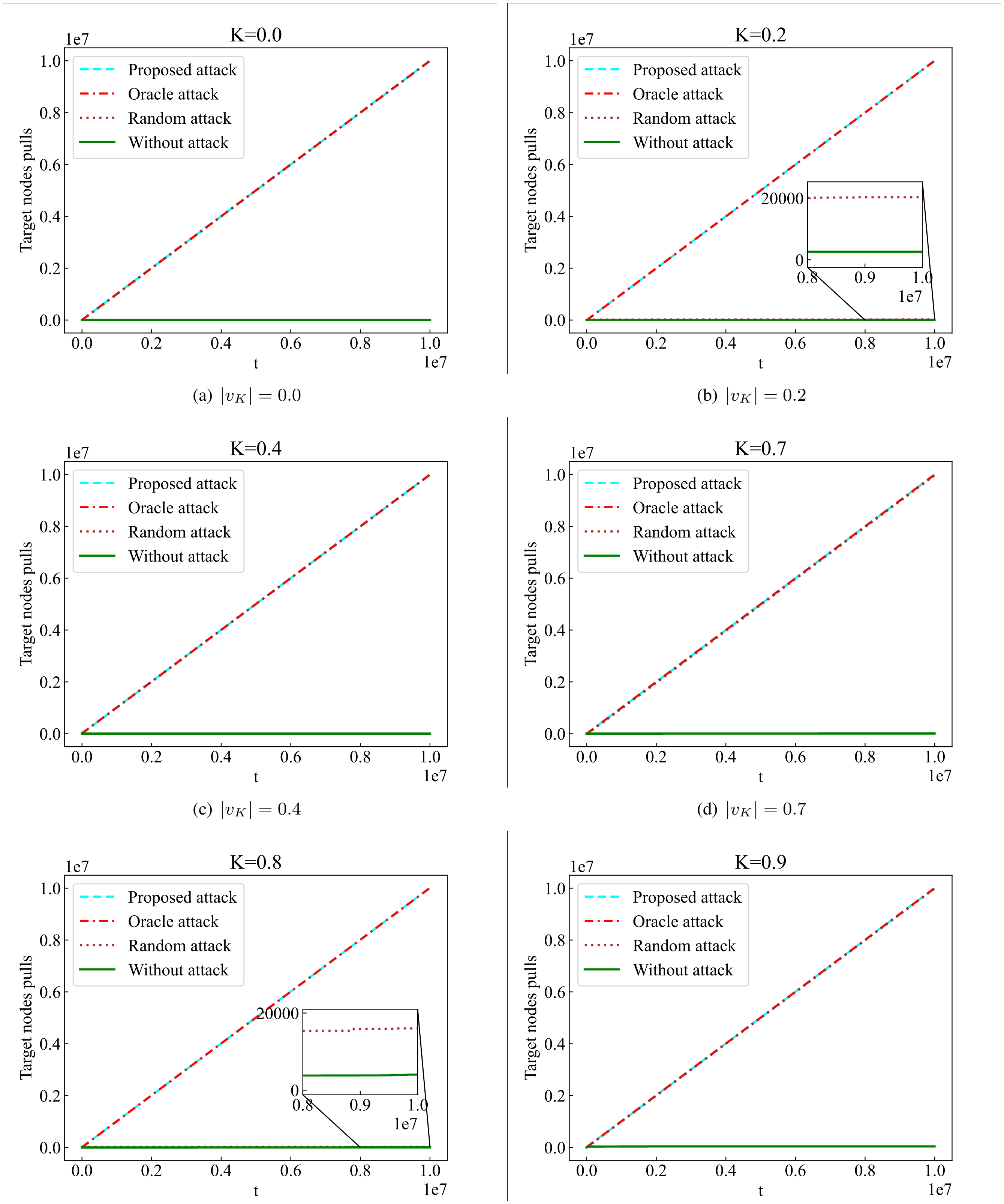
(c) $|v_K| = 0.4$

(d) $|v_K| = 0.7$

(e) $|v_K| = 0.8$

(f) $|v_K| = 0.9$

Figure 3: The graphs of $C(t) - t$



(a) $|v_K| = 0.0$

(b) $|v_K| = 0.2$

(c) $|v_K| = 0.4$

(d) $|v_K| = 0.7$

(e) $|v_K| = 0.8$

(f) $|v_K| = 0.9$

Figure 4: The graphs of Target nodes pulls$-t$