Supplementary Information For "An Indirect-Reciprocity Game Theoretic Framework for Device-to-Device Multicast"

Yuyang Chen[1], Biling Zhang[2], Min Wang[2], Tianyu Xu[3], and Zhu Han[4]
1. Department of Computer Science, University of California, Irvin, USA.
2. School of Network Education, Beijing University of Posts and Telecom. (BUPT), China.
3. School of Information and Communication Engineering, BUPT, China.
4. Department of Electrical and Computer Engineering, University of Houston, Houston, USA.
Email: yuyangc3@uci.edu, bilingzhang@bupt.edu.cn, wang_min@bupt.edu.cn,
bupt_isaac@163.com, zhan2@uh.edu

I. Proof of Lemma 1

Lemma 1: Given $\boldsymbol{x}$ and P_l, $Pr(l_1|l) < Pr(l_2|l), \forall l_1 > l_2$. In other word, $Pr(l'|l)$ is a decrease function of l'.

Proof: Let $k = \frac{P_{th} - \mu(d_{l'}|l)}{\sigma_d(d_{l'}|l)}$.

Since $\sigma_d^2 \propto d_{l'}^{-4}$ and $\mu = M\sqrt{P_l}\sigma_d^2\sigma_c^2\sigma_n^2$, we have $\mu \propto d_{l'}^{-4}$. This means that as $d_{l'}$ increases, the value of both σ_d^2 and μ decreases, and thus k increases.

Noticing that function $Q(k)$ is a decrease function of k, it is also a decrease function of $d_{l'}$. Therefore, $Pr(l'|l)$ is a decrease function of l', i.e., $Pr(l_1|l) < Pr(l_2|l), \forall l_1 > l_2$. ■

II. Proof of Theorem 1

Theorem 1: There exists a unique stationary reputation distribution of the whole population for any given optimal action rule, and the stationary reputation distribution is $\boldsymbol{P_t}$'s eigenvector with the corresponding eigenvalue one.

Proof: $\forall i \in \{0, 1, 2, ..., L-1\}$, the element $P_{t_{i,j}}$ in $\boldsymbol{P_t}$ has a property as follow

$$\sum_{j=0}^{L-1} P_{t_{i,j}} = 1. \tag{1}$$

Getting the characteristic polynomial of $\boldsymbol{P_t}$, we have

$$\begin{vmatrix} \lambda - P_{t_{0,0}} & -P_{t_{0,1}} & \cdots & -P_{t_{0,L-1}} \\ -P_{t_{0,0}} & \lambda - P_{t_{1,1}} & \cdots & -P_{t_{1,L-1}} \\ \vdots & \vdots & \ddots & \vdots \\ -P_{t_{L-1,0}} & -P_{t_{L-1,1}} & \cdots & \lambda - P_{t_{L-1,L-1}} \end{vmatrix} \tag{2}$$

Adding the elements from the 2nd to the L^{th} column to the first column, and using (1), we have

$$\begin{vmatrix} \lambda - 1 & P_{t_{0,1}} & \cdots & P_{t_{0,L-1}} \\ \lambda - 1 & \lambda - P_{t_{1,1}} & \cdots & P_{t_{1,L-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda - 1 & P_{t_{L-1,1}} & \cdots & \lambda - P_{t_{L-1,L-1}} \end{vmatrix} \tag{3}$$

Noticing that the elements in the first column of (3) are all $\lambda - 1$, we can conclude that $\lambda = 1$ is a root of the characteristic polynomial, i.e., 1 is an eigenvalue of $\boldsymbol{P_t}$. Hence, $\boldsymbol{P_t}$ and its transpose must have an eigenvalue of 1, and the eigenvector corresponding to this eigenvalue is the solution to (12) in [1]. ■

III. Proof of Theorem 2

Theorem 2: The optimal action in (17) of [1] is an ESS if the cost-to-gain ratio c_e/g satisfies $0 < c_e/g < \frac{(1-\lambda)\beta W G_{\min}}{\bar{\eta} P_{\max}}$, where $\bar{\eta} = \sum_{l=1}^{L} l\eta_l$, $W = w_b\bar{B}\bar{N} + w_e\bar{P}$, $G_{\min} = \min_{\rho'} \sum\limits_{m}(G_{L,m} - G_{f(\rho'),m})$, and $\rho' = \rho(l')$.

Proof: Given the SL s and the LDM where only the outermost l ($l \in \mathcal{L}$) rings are not distributed with RNs, $\forall i,j$, the optimal action should be $a^*_{i,j} = P_l$. With P_l, the HN can serve all the RNs. According to the social norm in (3) of [1], the HN is assigned an immediat reputation $\Omega_{s,L} = L$.

Then let $\bar{\eta} = \sum_{l=1}^{L} l\eta_l$ and $W = w_b\bar{B}\bar{N} + w_e\bar{P}$, we have

$$\hat{r}(a^*_{i,j}) = e_{\Omega_{s,L}} = e_L, \tag{4}$$

$$\boldsymbol{r}(a^*_{i,j}) = (1-\lambda)\sum_{l} e_{\Omega_{s,l}} G_{L,l} + \lambda e_i = (1-\lambda)G_{L,:} + \lambda e_i, \tag{5}$$

$$t(a^*_{i,j}) = \frac{\sum_{l=1}^{L} l[(1-\lambda)G_{L,l} + \lambda e_i]}{\bar{\eta}}, \tag{6}$$

and

$$\begin{aligned} &F_i(a^*, a^*) \qquad (7) \\ &= -sd_{BH} - c_eP_l + g\eta W \tfrac{\sum_{l=1}^{L} l[(1-\lambda)G_{L,l}+\lambda e_i]}{\bar{\eta}} + \delta \sum_k \sum_m r_k(a_{i,m})U_{k,m}q_m \\ &= -sd_{BH} - c_eP_l + g\eta W \tfrac{\sum_{l=1}^{L} l[(1-\lambda)G_{L,l}+\lambda e_i]}{\bar{\eta}} + \delta\lambda U_{i,:}\boldsymbol{q}^T + \delta(1-\lambda)\sum_m G_{L,m}U_{m,:}\boldsymbol{q}^T, \end{aligned}$$

where $\boldsymbol{q} = [q_1, ..., q_L]$ and $\boldsymbol{q}^T$ is the transpose of $\boldsymbol{q}$.

$\forall a_{i,j} \neq a^*_{i,j}$, if $a_{i,j} = P'_l$ and $P'_l > P_l$, we have $\rho(l') = \rho(l)$, and thus $\hat{r}(a_{i,j}) = \hat{r}(a^*_{i,j})$, $\boldsymbol{r}(a_{i,j}) = \boldsymbol{r}(a^*_{i,j})$ and $t(a_{i,j}) = t(a^*_{i,j})$. Since $c_e > 0$ and $c_e P'_l > c_e P_l$ is always hold, we have $F_i(a^*, a^*) > F_i(a^*, a^*)$.

If $a_{i,j} = P'_l$ and $P'_l < P_l$, denoting $\rho(l') = \rho'$, we have

$$\hat{r}(a_{i,j}) = e_{\Omega_{s,f(\rho')}} = e_{f(\rho')}, \tag{8}$$

$$\boldsymbol{r}(a_{i,j}) = (1-\lambda)\sum_l e_{\Omega_{s,l}} G_{f(\rho'),l} + \lambda e_i = (1-\lambda) G_{f(\rho'),:} + \lambda e_i, \tag{9}$$

$$t(a_{i,j}) = \frac{\sum_{l=1}^{L} l[(1-\lambda)G_{f(\rho'),l} + \lambda e_i]}{\bar{\eta}}, \tag{10}$$

and

$$\begin{aligned} &F_i(a_{i,j}, a^*) \\ &= -sd_{BH} - c_e P_{l'} + g\eta W \tfrac{\sum_{l=1}^{L} l[(1-\lambda)G_{f(\rho'),l} + \lambda e_i]}{\bar{\eta}} + \delta \sum_k \sum_m r_k(a_{i,m}) U_{k,m} q_m \\ &= -sd_{BH} - c_e P_{l'} + g\eta W \tfrac{\sum_{l=1}^{L} l[(1-\lambda)G_{f(\rho'),l} + \lambda e_i]}{\bar{\eta}} + \delta\lambda U_{i,:}\boldsymbol{q}^T + \delta(1-\lambda)\sum_m G_{f(\rho'),m} U_{m,:}\boldsymbol{q}^T. \end{aligned} \tag{11}$$

Combining (7) and (11), we have

$$\begin{aligned} &F_i(a^*, a^*) - F_i(a_{i,j}, a^*) \\ &= -c_e(P_l - P_{l'}) + g\eta W \tfrac{\sum_{l=1}^{L} l(1-\lambda)[G_{L,l} - G_{f(\rho'),l}]}{\bar{\eta}} + \delta(1-\lambda)\sum_m (G_{L,m} - G_{f(\rho'),m}) U_{m,:}\boldsymbol{q}^T \\ &> -c_e P_{\max} + g\eta W \tfrac{\sum_{l=1}^{L} l(1-\lambda) G_{\min}}{\bar{\eta}} \end{aligned} \tag{12}$$

where $G_{\min} = \min_{\rho'} \sum_m (G_{L,m} - G_{f(\rho'),m})$.

Then for $F_i(a^*, a^*) > F_i(a^*, a^*)$ to be hold, we should have

$$g/c_e > \frac{P_{\max}\bar{\eta}}{(1-\lambda) W G_{\min}}. \tag{13}$$

■

IV. Algorithm 1

Algorithm 1 : Find the Optimal Action Using Value Iteration

1. Given the tolerance $\epsilon = 0.01$, and set $\epsilon_1 = 1$, and initialize $\boldsymbol{A}^* = \boldsymbol{A}^0$
2. While $\epsilon_1 > \epsilon$
   - Set $\epsilon_2 = 1$
   - For $i = 0 : 9$ and $j = 1 : 10$
   - Initialize $U_{i,j} = 0, \forall i, \forall j$
   - While $\epsilon_2 > \epsilon$
     - Compute reputation distribution $\boldsymbol{r}$.
     - Find the stationary reputation distribution η^*.
     - Calculate the number of vacant channels $t(a_{i,j})$.
     - Obtain $\widehat{U}_{i,j}$ based on the stationary reputation distribution $\hat{\boldsymbol{\eta}}^*$.
     - Find Find the optimal action $\widehat{a}^*_{i,j} = arg \max_{a_{i,j}} \widehat{U}_{i,j}$.
     - Update the parameter$\epsilon_2 = (U_{i,j} - \widehat{U}_{i,j})^2$.
     - Update $U_{i,j}$ with $U_{i,j} = \widehat{U}_{i,j}$
     - End
   - End While
   - End For
   - Update the parameter$\epsilon_1 = ||A^* - \widehat{A}^*||^2$.
   - Update A^* with $A^* = \widehat{A}^*$

   End While

References

[1] Y. Chen, B. Zhang, M. Wang, T. Xu, and Z. Han,“An Indirect-Reciprocity Game Theoretic Framework for Device-to-Device Multicast”, submitted to ICC 2019.