Online Appendix for "Opportunistic Information Networks"

Ahmed M. Alaa, *Member, IEEE*, Kartik Ahuja, and Mihaela van der Schaar, *Fellow, IEEE*

Appendix A
Proof of Theorem 1

From *Nash's Existence Theorem*, we know that if we allow mixed strategies, then every game with a finite number of players in which each player can choose from finitely many pure strategies has at least one Nash equilibrium [28]. Assume that agent i adopts a mixed strategy $\Delta_i = (p_{i1}, p_{i2}, ..., p_{iN})$, where p_{ij} is the probability that agent i forms a link with agent j, and $p_{ii} = 0, \forall i \in \mathcal{N}$. The utility of agent i in this case is obtained by averaging over all possible networks as follows

$$u_i(\Delta_i) = \sum_{j=1}^{2^{N-1}-1} w_j f\left(H(X_i \cup X_{\alpha_j})\right) - \sum_{l=1}^{N-1} p_{il} c, \tag{A.1}$$

where α_j is an element of the power set of $\mathcal{N}/\{i\}$, and w_j is the probability of the emergence of a network component comprising agents in the set $\{i \cup \alpha_j\}$ based on the mixed strategies. For instance, in a 2 agent network, the utility function of agent 1 is given by

$$u_1(\Delta_1) = \left(p_{12}(1-p_{21}) + p_{21}(1-p_{12}) + p_{12}p_{21}\right) f\left(H(X_1, X_2)\right)$$

$$+(1-p_{12})(1-p_{21}) f\left(H(X_1)\right) - p_{12}c.$$

In this case, $w_1 = p_{12}(1-p_{21}) + p_{21}(1-p_{12}) + p_{12}p_{21}$ and $w_2 = (1-p_{12})(1-p_{21})$. Let the NE strategy profile be $\boldsymbol{\Delta}^* = (\Delta_1^*, \Delta_2^*, ..., \Delta_N^*)$, where $\Delta_i^* = (p_{i1}^*, p_{i2}^*, ..., p_{iN}^*)$. According to (3), the following condition on $\boldsymbol{\Delta}^*$ needs to be satisfied

$$u_i(\Delta_i^*, \Delta_{-i}^*) \geq u_i(\Delta_i, \Delta_{-i}^*), \forall \Delta_i \in [0,1]^N, \forall i \in \mathcal{N}. \tag{A.2}$$

Now we show that for any agent i, the NE strategy Δ_i^* needs to be a pure strategy for condition (A.2) to be satisfied. We focus on agent i with a NE strategy $\Delta_i^* = (p_{i1}^*, p_{i2}^*, ..., p_{iN}^*)$, where $p_{ij}^* \in [0,1]$. Now assume we induce a perturbation ϵ to the mixed strategy of agent i by modifying p_{ik}^* to $p_{ik}^* + \epsilon$ for a certain k, where $\epsilon \in [-p_{ik}^*, 1-p_{ik}^*]$. We call this modified strategy $\Delta_i^*(\epsilon)$. Note that we can write any w_j in (A.1) in the form of $w_j = \tilde{w}_j p_{ik}^* + \bar{w}_j(1-p_{ik}^*)$. This results in a perturbed utility $u_i(\Delta_i^*(\epsilon))$ as follows

$$u_i(\Delta_i^*(\epsilon)) = \sum_{j=1}^{2^{N-1}-1} \left(\tilde{w}_j (p_{ik}^* + \epsilon) f\left(H(X_i \cup X_{\alpha_j})\right)\right.$$

The authors are with the Department of Electrical Engineering, University of California Los Angeles (UCLA), Los Angeles, CA, 90095, USA (e-mail: ahmedmalaa@ucla.edu, ahujak@ucla.edu, mihaela@ee.ucla.edu). This work was funded by the Office of Naval Research (ONR).

$$+\bar{w}_j(1-\epsilon-p^*_{ik})f\left(H(X_i\cup X_{\alpha_j})\right)\Big) - (p^*_{ik}+\epsilon)c - \sum_{l=1,l\neq k}^{N-1} p^*_{il}c, \tag{A.3}$$

which can be rearranged as

$$u_i(\Delta_i^*(\epsilon)) =$$
$$u_i(\Delta_i^*) + \epsilon\left(\sum_{j=1}^{2^{N-1}-1}(\tilde{w}_j-\bar{w}_j)f\left(H(X_i\cup X_{\alpha_j})\right) - c\right).$$

Let $\delta = \sum_{j=1}^{2^{N-1}-1}(\tilde{w}_j-\bar{w}_j)f\left(H(X_i\cup X_{\alpha_j})\right) - c$. It can be easily shown that $\frac{\partial u_i(\Delta_i^*(\epsilon))}{\partial \epsilon} > 0$ if $\delta > 0$, and $\frac{\partial u_i(\Delta_i^*(\epsilon))}{\partial \epsilon} < 0$ otherwise. Thus, if $\delta > 0$, agent i can always increase its utility by increasing ϵ and setting $\epsilon = 1 - p^*_{ik}$ (and thus playing a pure strategy with $p_{ik} = 1$), and if $\delta < 0$, agent i can always increase its utility by setting $\epsilon = -p^*_{ik}$(and thus playing a pure strategy with $p_{ik} = 0$), which contradicts with Δ_i^* being a NE strategy. Thus, for all $k \in \mathcal{N}/\{i\}$, agent i needs to select a pure strategy $p^*_{ik} \in \{0,1\}$ for Δ_i^* to be a best response to Δ^*_{-i} regardless of the strategies of other agents, i.e. non-pure strategies are always dominated by a pure strategy. Due to symmetry, this applies to all agents in $\mathcal{N}$. Therefore, it follows that a pure strategy NE always exists.

Appendix B
Proof of Proposition 1

If the component $\mathcal{C}$ is not minimally connected, then it has at least one cycle as there exist agents i and j that are connected via two paths $p_{ij,1}$ and $p_{ij,2}$, such that any of the two paths is not a subset of the other. For such component at NE, assume that agent v is on path $p_{ij,1}$ and agent w is on path $p_{ij,2}$. Note that all the agents receive the same amount of total information $H(\mathcal{C})$. We know that there indeed exists links: g^*_{xv} (or g^*_{vx}) and g^*_{wy} (or g^*_{yw}), where agent $x \in p_{ij,1}$ and agent $y \in p_{ij,2}$. Now focus on any link of them, say $g^*_{wy} = 1$. We observe that agent w can break this link and still receive the same benefit by gathering the same amount of information from path $p_{ij,1}$, thus receiving a strictly higher utility function as it will not pay the cost for the link with agent y, which contradicts the fact that $\mathbf{g}^*$ is an NE. Thus, a single path exists between any two agents.

Appendix C
Proof of Lemma 1

If there exists an agent in which other agents have an incentive to connect to even if they possess all other information in the network, then the network is indeed connected at any equilibrium. This is satisfied if and only if the linking cost satisfies $c < f(H(\mathcal{X})) - f(H(X_i))$ for some agent i in $\mathcal{N}$, i.e. the marginal benefit from connecting to that agent is always more than the link cost irrespective to the current connections of the agent forming the link. Thus, we must have $c < \max_i f(H(\mathcal{X})) - f(H(X_i))$. Hence, part (i) of the Lemma follows.

If no agent have an incentive to form any link, then the network is fully disconnected. From the monotonicity property of the entropy, we know that if agent i has no incentive to connect to a set $\mathcal{V}$ of agents, then it has no incentive to connect to a set $\mathcal{U}$ if $\mathcal{U} \subseteq \mathcal{V}$. Thus, if agent i has no incentive to connect to the set $\mathcal{N}/\{i\}$ via a single link, then it has no incentive to form any link in the network. This occurs if $c > f(H(\mathcal{X})) - f(H(X_i))$. If this condition is satisfied for all agents, then the network is indeed disconnected, and part (ii) of the Lemma follows.

APPENDIX D
PROOF OF THEOREM 2

For the network to be in NE, no agent should have an incentive to unilaterally deviate by forming a new link or breaking a link. Focus on a certain component $\mathcal{C}_i$. Inside this component, each agent should either have an incentive to form at least one link, or other agents should have an incentive to connect to it. Otherwise, this agent can be removed from the component while strictly increasing the utility of some agent. Thus, we must have either $f(H(X_{\mathcal{C}_i})) - f(H(X_j)) > c$ or $f(H(X_{\mathcal{C}_i})) - f(H(X_{\mathcal{C}_i/\{j\}})) > c$ for all agents j in $\mathcal{C}_i$. This should apply to all components in the network. Hence, condition (1) follows.

Now focus on the interaction between different components of the network. If any agent in component $\mathcal{C}_i$ benefits from forming a link to any agent in component $\mathcal{C}_j$, then the network is not NE since in this case an agent in $\mathcal{C}_i$ can strictly increase its utility by unilateral deviation. Hence, we should have $f(H(X_{\mathcal{C}_i \cup \mathcal{C}_j})) - f(H(X_{\mathcal{C}_i})) \leq c$ for any two components in the network. Thus, condition (2) follows.

APPENDIX E
PROOF OF LEMMA 2

We know that in the $\mathcal{K}_C$ region, all the NE networks are connected. Thus, the social welfare of any network in $\mathcal{K}_C$ is given by $U(\mathbf{g}^*) = Nf(H(\mathcal{X})) - (N-1)c$. The socially optimal network in $\mathcal{K}_C$ is the one with a social welfare of $\tilde{U}$, where $\tilde{U} = \max_{\mathbf{g} \in \mathbf{G}} U(g)$. Since $f(H(\mathcal{X})) - f(H(X_i)) > c, \forall i$ in the $\mathcal{K}_C$ region, then it is clear that a connected network maximizes the social welfare. Therefore, $\tilde{U}$ = $U(\mathbf{g}^*)$ and the PoA = 1 in the $\mathcal{K}_C$ region. Next, we focus on the $\mathcal{K}_I$ region. In this region, any connection will result a negative payoff for any agent who forms a link since $c > f(H(\mathcal{X})) - f(\min_i H(X_i))$. Thus, the social optimal is a fully disconnected network, which is also the unique (strict) NE, and the PoA = 1 in the $\mathcal{K}_I$ region. For the $\mathcal{K}_M$ region, the maximum PoA will occur if a fully disconnected network is an equilibrium and a connected network is a social optimum. In what follows, we show that this is indeed possible. Consider the case when $f(H(X_i, X_j)) - f(H(X_i)) < c, \forall i, j$, and $f(H(\mathcal{X})) - f(H(X_i)) > c, \forall i$. In this case, agents do not get immediate benefit from forming links to individual agents, thus a fully disconnected network is an NE since not forming a link is a best response for all agents when all other agents do not form a link. Therefore, the PoA in the $\mathcal{K}_M$ region is upper bounded by the social welfare of a connected network and that of a fully disconnected network, i.e. PoA $\leq \frac{Nf(H(\mathcal{X}))-(N-1)c}{\sum_{i=1}^N f(H(X_i))}$.

APPENDIX F
PROOF OF THEOREM 4

The PoA can be written as PoA $\leq \frac{Nf(\sum_{i=1}^N H(X_i) - D(\mathbf{p}||\mathbf{q})) - (N-1)c}{\sum_{i=1}^N f(H(X_i))}$. Note that the benefit function $f(x)$ is monotonically increasing in x. Thus, as $D(\mathbf{p}||\mathbf{q})$ increases, $f(\sum_{i=1}^N H(X_i) - D(\mathbf{p}||\mathbf{q}))$ decreases, and the PoA decreases consequently. Therefore, we have $\frac{\partial \text{PoA}}{\partial D(\mathbf{p}||\mathbf{q})} < 0$.

APPENDIX G
PROOF OF COROLLARY 1

In the $\mathcal{K}_C$ region, we know that all NE networks are connected. Thus, $\sup_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = \inf_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = H(\mathcal{X})$, and MIL $= 0$. Similarly, in the $\mathcal{K}_I$ region, we have $\sup_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = \inf_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = \min_i H(X_i)$, thus MIL $= 0$. In the $\mathcal{K}_M$ region, the MIL is maximized if both a connected and a fully disconnected network are equilibria. In this case, $\sup_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = H(\mathcal{X})$, and $\inf_{\mathbf{g}_u^* \in G^*} H(X_i \cup X_{\mathcal{R}_i(\mathbf{g}_u^*)}) = \min_i H(X_i)$. Thus, MIL $\leq H(\mathcal{X}) - \min_i H(X_i)$.

APPENDIX H
PROOF OF COROLLARY 2

From Theorem 5, we know that when $c > k\bar{H}$, then we have a unique equilibrium $\mathbf{s}^*$ for both $\tilde{F}_{\mathcal{H}}$ and $\mathcal{L}_{\mathcal{H}}$ in which $g_{ij}^* = 0, \forall i, j \in \mathcal{N}$, and $H^*(X_i) = \bar{H}$. Thus, we have $H^*(X_i) > 0, \forall i \in \mathcal{N}$, and $\frac{|\mathcal{I}(\mathbf{s}^*)|}{N} = 1$, which applies when the number of agents in the CIN grows to infinity, hence (9) follows. Next, we focus on the total amount of information in the network. For $\mathcal{L}_{\mathcal{H}}$, we have $H(X_1, X_2, ..., X_N) = \max\{\bar{H}, \bar{H}, ..., \bar{H}\} = \bar{H}$, and (10) follows. Finally, for $\tilde{F}_{\mathcal{H}}$, we have $H(X_1, X_2, ..., X_N) = \sum_{i=1}^{N} \bar{H} = N\bar{H}$, and (11) follows.

APPENDIX I
PROOF OF COROLLARY 3

We start by deriving (12). From Theorem 6, we know that for $\tilde{F}_{\mathcal{H}}$, every equilibrium has only one information producer. When the number of agents grows to infinity, we will still have one information producer and $\frac{|\mathcal{I}(\mathbf{s}^*)|}{N} = 0$. In order to prove (13), one needs to find one network in equilibrium for $\tilde{F}_{\mathcal{H}}$ in which, for arbitrary N, we have N information producers. Consider this network for N agents. Assume that $H(X_i) = \frac{\bar{H}}{N}, \forall i \in \mathcal{N}$, and the network has a single component which is periphery-sponsored star network. It is clear that for this network, $|\mathcal{I}(\mathbf{s})| = N$. We want to show that this network is an NE by showing that every agents strategy is best response to all others. It is easy to see that since $c < k\bar{H}$, each periphery agent has no incentive to break its link with the core since $\frac{N-1}{N} k\bar{H} > c$ when N is asymptotically large. Moreover, no agent has incentive to alter its information production profile since the total information in the network is $\sum_{i=1}^{N} \frac{\bar{H}}{N} = \bar{H}$. Thus, $\mathbf{s}$ is an NE. Since this applies to any N, (13) follows. Finally, since the network is always connected in any equilibrium, then (14) directly follows.

REFERENCES

[1] K.-C. Chen, M. Chiang, and H. V. Poor, "From Technological Networks to Social Networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 548-572, Sep. 2013.

[2] K. C.-J. Lin, C.-W. Chen, and C.-F. Chou, "Preference-Aware Content Dissemination in Opportunistic Mobile Social Networks," *in Proc. of IEEE INFOCOM 2012*.

[3] O. Yagan, D. Qian, J. Zhang, D. Cochran, "Conjoining speeds up information diffusion in overlaying social-physical networks", *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1038-048, 2013.

[4] W. Saad, Z. Han, A. Hjorungnes, D. Niyato, E. Hossain, "Coalition Formation Games for Distributed Cooperation Among Roadside Units in Vehicular Networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 48-60, Jan. 2011.

[5] A. P. Athreya and P. Tague, "Network Self-Organization in the Internet of Things," *IEEE International Workshop on Internet-of-Things Networking and Control*, 2013.

[6] P. K. Biswas, and S. Phoha, "Self-Organizing Sensor Networks for Integrated Target Surveillance," *IEEE Trans. Computers*, vol. 55, no. 8, pp. 1033-1047, Aug. 2006.

[7] H. Sugiyama, T. Tsujioka, and M. Murata, "Autonomous chain network formation by multi-robot rescue system with ad hoc networking," *2010 IEEE International Workshop on Safety Security and Rescue Robotics (SSRR),* Bremen, 26-30 July, 2010.

[8] M. O. Jackson, "Social and Economic Networks," *Princeton University Press*, 2008.

[9] N. Eagle, A. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 106, no. 36, pp. 274278, 2009.

[10] M. ElSherief, T. ElBatt, A. Zahran, and A. Helmy, "The quest for user similarity in mobile societies," *in Proc. of the 2nd SCI Workshop in conjunction with IEEE PerCom 2014*, 2014.

[11] A.-K. Pietiladnen and C. Diot, "Dissemination in opportunistic social networks: the role of temporal communities," *in Proc. of ACM Mobihoc*, 2012.

[12] V. Bala and S. Goyal,"A non-cooperative model of network formation," *Econometrica*, vol. 68, no. 5, pp. 1181-1231.

[13] M. Jackson and A. Watts, "On the Formation of Interaction Networks in Social Coordination Games,"*Games and Economic Behavior*, vol. 41, pp. 265-291, 2002.

[14] Y. Zhang and M. van der Schaar, "Information Production and Link Formation in Social Computing Systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2136-2145, Dec. 2012.

[15] Y. Zhang and M. van der Schaar, "Strategic Networks: Information Dissemination and Link Formation Among Self-Interested Agents," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1115-1123, Jun. 2013.

[16] J. Xu, S. Zhang, and M. van der Schaar, "Network Dynamics with Incomplete Information and Learning," *Allerton* 2014.

[17] Y. Song and M. van der Schaar, "Dynamic Network Formation with Incomplete Information," accepted and to appear in *Economic Theory*, 2015.

[18] J. Barros and S. D. Servetto, "Network Information Flow With Correlated Sources," *IEEE Trans. Inform. Theory,* vol. 52, no. 1, Jan. 2006, pp. 155-170.

[19] Y. Sung, H. V. Poor, and H. Yu, "How Much Information Can One Get From a Wireless Ad Hoc Sensor Network Over a Correlated Random Field?," *IEEE Trans. Inform. Theory,* vol. 55, no. 6, Jun. 2009, pp. 2827-2847.

[20] H. Gupta, V. Navda, S. R. Das, V. Chowdhary, "Efficient Gathering of Correlated Data in Sensor Networks," *ACM Transactions on Sensor Networks (TOSN),* vol. 4, no. 1, Jan. 2008.

[21] A. Giridhar and P. R. Kumar, "Computing and Communicating Functions Over Sensor Networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 755-764, Apr. 2005.

[22] H. Luo, Y. Liu, and S. K. Das, "Routing Correlated Data with Fusion Cost in Wireless Sensor Networks," *IEEE Trans. Mobile Comp.* , vol. 5, no. 11, pp. 1620-1632, Nov. 2006.

[23] A. Ramamoorthy, "Minimum Cost Distributed Source Coding Over a Network," *IEEE Trans. Inform. Theory,* vol. 57, no. 11, Jan. 2011.

[24] A. Lee, M. Medard, K. Z. Haigh, S. Gowan, and P. Rubel, "Minimum-Cost Subgraphs for Joint Distributed Source and Network Coding," *Proc. of Workshop on Network Coding, Theory and Applications*, 2007.

[25] A. Galeotti and S. Goyal, "The Law of the Few," *American Economic Review,* vol. 100, no. 4, pp. 1468-1492, 2010.

[26] Z. Zhang and R. W. Yeung, "On Characterization of Entropy Function via Information Inequalities," *IEEE Trans. Inform. Theory,* vol. 44, no. 4, pp. 1440-1452, Jul. 1998.

[27] T. M. Cover and J. A. Thomas, "Elements of Information Theory," *New York, Wiley*, 2006.

[28] J. Nash,"Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 36, no. 1, pp. 48-49, 1950.