# Discovering Trustworthy Social Spaces

Udayan Kumar Ahmed Helmy
{ukumar,helmy}@cise.ufl.edu
University of Florida, Gainesville, USA

## ABSTRACT

Many future mobile services and applications will center on the social and community aspects of mobile societies. Interactions and connections between users in mobile networks are usually subject to the strength of the connections between the nodes, informed by historical events. This study proposes, implements and evaluates novel methods to dynamically measure the strength of social connections and similarity based on historical mobility behavior and encounter information. Through our protocol and application we investigate the feasibility of discovering known encountered devices, in addition to the opportunistic identification of potentially-strong new connections.

We propose a set of 4 filters to rate and rank mobile encounters identifying users with similar behavior. We have developed and deployed ConnectEnc application on Android and Nokia N810 platform to measure the link between the scores of proposed filters and the existence (or lack) of social relationship with the rated devices. We find that a statistically strong relationship exists between our recommendation and social relationship with the devices rated by the users (for LVC, r=0.84, p <0.01).

With this similarity based trustworthy node discovery, several potential applications can be enabled including mobile social networking, building groups and communities of interest, localized alert and emergency notification, context-aware and similarity-based networking.

## Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless Communication; C.2.m [**Miscellaneous**]: [opportunistic connections, Trust Framework, performance]

## Keywords

Trust Framework, Mobile Networks

## 1. INTRODUCTION

Most interactions between people consider, and in many cases, rely on prior established trust. In interactions over computer networks, trust establishment is particularly important and challenging, given the identity theft, fraud and security breaches common in the Internet today. These problems are further exacerbated by the uncertainty and dynamics in wireless mobile networks. Furthermore, in infrastructureless peer-to-peer mobile networks, such as Ad-hoc, delay tolerant (DTNs) or sensor networks, cooperation, and subsequently trust is imperative to the construction and operation of the network. With trust, several potential applications can be enabled including mobile social networking, groups and communities of interest, localized alert and emergency notification, context-aware and similarity-based networking [1], to name a few.

In this work, we introduce *ConnectEnc*, a mobile framework, that utilizes mobile encounters to augment user's view of surrounding social space. This application can be helpful in re-linking with already known social connections or discovering new social similar connections. *ConnectEnc* application draws inspiration from the social-science principle of Homophily [2],which states that people with similar interest meet and interact often. Discovering frequently encountered users would mean discovering people similar to oneself (like work colleagues or classmates). Encounters based on mobility provide opportunities to build proximity, location and similarity based trust. Each mobile device running our application keeps simple history of other devices encountered.

An encounter, in our case, can be defined as an event when two devices can detect radio signals from each other. *ConnectEnc* currently uses Bluetooth radio to measure encounters. When having an encounter the users are within the radio range (for Bluetooth range is ∼15m). This physical proximity gives the users an opportunity to meet face-to-face with each other (also verify identity) and exchange out-of-band information such cryptographically strong keys [3]. These out-of-band exchanged keys can now allow encountering users to have secure and authentic communication over any mutually agreed medium. This feature of trustworthy out-of-band information exchange is not possible on wired networks (two terminals may be geographically apart) but can be easily leveraged by *ConnectEnc*.

We propose four new filter to measure encounter based trust. We have analyzed filters using real world traces (more details in [4]) and found the filter scores to be stable over time and have low correlation among each other, implying encounters are evaluated on rich set of parameters thus the recommendations can be customized for each user. The four filters *ConnectEnc* application features includes, (1) Frequency of Encounters (FE) - total count of encounters, (2) Duration of Encounter (DE) - total duration of encounters, (3) Location Vector - Count (LV-C) and (4) Location Vector


*PhoneSense'12,* November 6, 2012, Toronto, ON, Canada.

- Duration (LV-D). The last two filters use spatio-temporal encounter data (combine location information with encounters) when measuring count and duration, thus providing location based behavioral similarity/homophily. Using these filters, *ConnectEnc* application provides the user with an option to choose trustworthy nodes in coordination with personal preferences, location priorities and/or contextual information.

We also conducted a user study, where participants carried *ConnectEnc* running devices for at least a month. Users were asked to mark the users they trust and would be willing to communicate with. We find usersâĂŹ willingness to communicate with other devices to be highly correlated with behavioral similarity. User feedback from using ConnectEnc app, shows that statistically strong correlation exists between the similarity scores generated by the filters and the selection of devices by the user. Proposed filters are able to capture 80% of the already known user within top 25% of the encountered users. Thus ConnectEnc is able to capture socially similar users with good success.

Overall, *ConnectEnc* augments the user view of the neighboring devices and keeps history of encountered devices, frequency of encounter and location along with other statistics. It also provides recommendations about potential trustworthy devices among those encountered based on a combined filter that our research group has introduced. A user may decide to meet a device's owner face-to-face to establish a trusted relationship for future communications. An accompanying registration service (optional) enables users to map device identifications (using mac addresses) to actual user names.

Currently, *ConnectEnc* is capable of running on Android platform and Linux based devices. In the following section, we present the application architecture.

## 2. RELATED WORK

Mobile devices are now becoming our constant companions. Traveling to places wherever the user visits, connecting the user to other users, internet and various information sources. The strength of link between user and the userâĂŹs mobile devices is ever stronger, so much so that they have become a user's alter ego. The presence of several powerful sensors, processing power and decent battery life can be used to convert mobile device into sensors of modern society. Researchers have invented several techniques to sense the activities of user [5]. A comprehensive survey of sensing techniques can be found here [6]. Emotion sensing techniques are also being developed [7] and also a large number of participatory sensing applications have also been developed [8].

We, however, find that not much work has been done in sensing relationships between two users of mobile devices. The presence of relationship information can not only add a sense of trust and security in existing applications such as disaster relief and peer-to-peer networking but can also lead to the creation of new services such as interest-based networking among others. The physical proximity between the two users can lead to better verification of user identity (via face-to-face interactions) and establishment of out-of-band encryption keys[9, 3, 10].

This trust information can also be utilized to bootstrap reputation and recommendations systems. In [11, 12, 13] there is an assumption of preexisting trust when the system starts, the proposed *ConnectEnc* framework can generate and provide trust information and thus bootstrap reputation and recommendations systems.

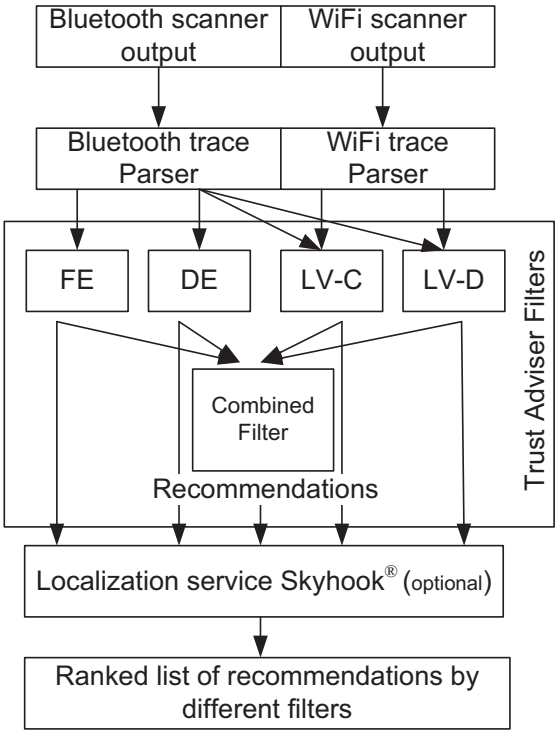



**Figure 1: Block diagram of the application**

In this current work, we draw inspiration from social science research [2] [14] to design algorithms/filters to capture and utilize relationship information based on behavioral similarity between the users to generate . To the best of our knowledge, this is the first work in this research area.

## 3. APPLICATION BLOCK DIAGRAM

Application block diagram is shown in Fig. 1. The arrows in the diagram represent how the encounter data flows in the application. The basic blocks of *ConnectEnc* are Bluetooth and Wi-Fi Scanning. Bluetooth scanning is used to discover and record Bluetooth devices and Wi-Fi scanning is used to obtain localization information. Traces from both the scanners are then parsed and given to the Trust Adviser Filters (Sec. 4). Encounters are then rated and ranked by filters. User can also choose to update locations which entails going to third party server such as Google and Skyhook to get location data based on the Wi-Fi AP data (users can also switch to more power hungry GPS for localization). Once the processing is done for a set of encounter data, the user are sorted (depending on the sorting key selected) and presented to the user. For the Combined Filter (details below), user can also select the weights.

In the following section, we go into the details of the Trust Adviser Filters.

## 4. TRUST ADVISER FILTERS

When users encounter; i.e., are within the radio range of each other (∼15m for Bluetooth), they can potentially exchange out-of-band information including identity information and cryptographic keys [3]. Such exchange is not

possible in wired networks. The function of providing meaningful, stable scores for encountered devices lies within the Trust Adviser Filters, which is the heart of ConnectEnc. In the implementation, a user would select users to trust (or interact) and the filters would serve as an adviser. Thus, users would have full control over the selection of trusted users. These filters would act as the scoring system that recommends users who are most similar to the user. We investigate five filters, based on: *i.* Simple encounter (frequency and duration) ranking, *ii.* Behavior similarity, and *iii.* Combined filter that combines the weighted scores of the other filters

## 4.1 Simple Encounter Ranking

These filters measure similarity by aggregating the encounter data using simple statistics. We present two such filters:

**Frequency of Encounters (*FE*):** ranks encountered devices based on total number of encounters over a window of history, regardless of the duration.

**Duration of Encounters (*DE*):** ranks encountered devices based on the total duration of encounters.

## 4.2 Behavior Similarity

Behavior based similarity measures similarity based on location visitations and preferences. We couple location information with encounters to determine the similarity between users.

**Location Vector (*LV*):**

To capture behavioral characteristic, we have designed a filter that stores location preferences of a user in a single dimensional vector. It is assumed for this filter that a device has some localization capability, which is quite common for today's devices. Every device maintains a vector for itself and a vector for each of the users it encounters. The columns of the vectors represent the different locations visited by a user and the values stored in each cell indicate either duration ($LV-D$) or count ($LV-C$) of the sessions at that particular location. For every encounter, the vector for the encounter node is updated with respect to the encounter location.

To get similarity score with another user, the inner product of one's location vector and the vector the user has maintained for other users is computed. The inner product gives a score indicating the similarity between two users. This score is higher if the two $LV$s are similar and can be zero, if the users do not have any visited location in common. Fig. 2 illustrates the design of a *location vector*. The idea of maintaining location vectors for each user allows *ConnectEnc* to find similarity score without any exchange of information between the users, helping to preserve users' privacy.

Since vectors for all the encountering users are maintained locally on the device, LV requires no exchange of vectors among users for calculating similarity. This is more privacy-preserving and more resilient to attacks since only first-hand information is used (equivalent to what user might have observed). This privacy comes at the cost of requiring extra storage space for storing vectors for each user. Considerable storage optimization is achieved by storing (for each encountering user) only the locations where encounters happened.

## 4.3 Combined Filter (*CF*)

Each filter provides a different perspective on an encounter or behavioral aspect. The Combined filter provides a systematic and flexible mechanism to combine the scores from all filters and present a unified score to the users. The selection of weights for various filters would depend on several factors including user's preference and feedback (check Sec. 5) and application requirements. A generic CF score ($C$) for a user $U_j$ can be generated by using the following:

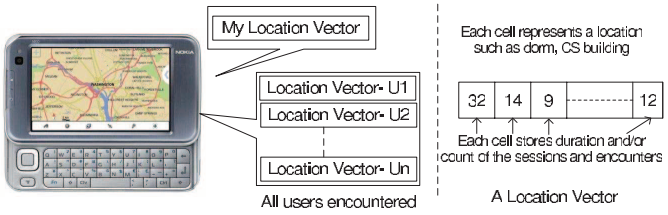



**Figure 2: Location Vector $LV$ for a user**

$$C(U_j) = \sum_{i}^{n} \alpha_i F_i(Uj) \quad (1)$$

where $F_i(U_j)$ is the normalized score for user $U_j$ according to filter $i$. The $\alpha_i$ is the weight given to filter score $F_i$ and $n$ is the total number of filters used. We select $\alpha_i$ such that $\sum \alpha_i = 1$, and $0 \leq \alpha_i \leq 1$.

This linear combination is chosen for its simplicity[1]. Our implementation allows users to customize these weights. From the analysis of user feedback (Sec. 5), we find that not all the users prefer same weights.

**Decay of filter scores:** Social science studies have shown that social relationship are dynamic and require frequent interactions to prevent decay. The strength of relationship wanes with the increase in time between interactions. This decay follows a exponential decay pattern with half time dependent on the relationship type [14] (3.5 years for family, 6 months for colleagues). Configurable decay was integrated in our *ConnectEnc* app with default halftime set to 6 months.

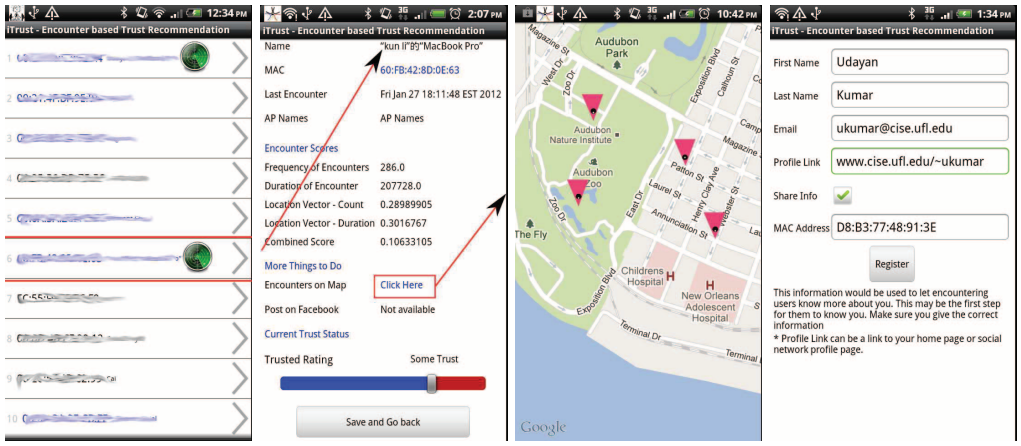
# 5. IMPLEMENTATION AND VALIDATION

Currently, *ConnectEnc* is available for Android platform and Linux based Nokia Tablet N810. It provides the ability to rate encounter users based on FE, DE, LV and CF. Encountered users can be sorted by any filter and weights for the CF are user configurable. If some of the encountered users are currently discoverable, their listing would have a green circular mark as shown in Fig. 4A.. The application provides inbuilt facilities for scanning Bluetooth devices and Wireless Access Points (for localization as GPS is energy-wise expensive. User can select GPS, if needed).

On selecting a particular user, encounter details (Fig. 4B. are presented and clicking on the map option one can see encounter locations on map (Fig. 4C.). Encountering devices can be rated for trust by the user on the scale from -2 (no Trust) to 2 (high Trust). This allows users to store their evaluations for the encountered devices and can be also used by other applications on the user's device.

One of the challenges faced by mobile ad-hoc communications or services such as *ConnectEnc*, in general, is the unknown identity of the other encountering device. In *ConnectEnc*, we solve this problem by incorporating a registry database lookup where information about a registered user can be obtained based on the MAC address of the device

[1]Other non-linear combinations shall be investigated in future work.



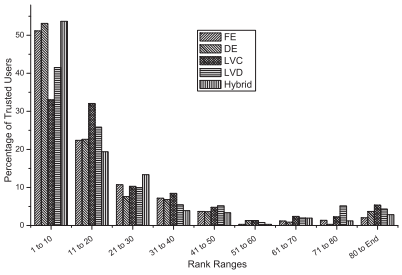

A. Landing page with current encounter marked with Green circles

B. Details for an encountered user

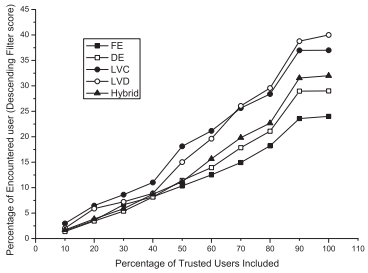
C. Encounters on Map

D. Registration Screen

**Figure 4: Screenshots of *ConnectEnc* application**




A.




B.

**Figure 3: *ConnectEnc* evaluations based on application usage. Fig. A shows the percentage of trusted users in 1 to 10 Top user, 11 to 20 Top users for each filter. Fig B. shows fraction of encounter users needed (from top) to capture 'x'% of trusted users for each filter**

(Users can opt-out of database registration and can also make information private such that their consent is required before any info is shared) 4D.. By maintaining a registry of devices and owners, encountering devices can get more information before having face to face encounters. Using *ConnectEnc*, one can also setup up alerts if a particular device is in the encounter range. Alerts can allow users to send message to other user when they are in radio range.

**Application Evaluation:** Initially we had 30 CS major volunteer (grad and undergrad) running *ConnectEnc*, out of which only 22 users ran *ConnectEnc* consistently for at least a month. For the evaluations, we have only used traces from these 22 users. The users were asked to mark devices they trust. On average, number of trusted user marked by each user is 15 and number of unique devices encountered per user is 175. We use this data to investigate if behavioral similarity as captured by the trust filters correlates to trusted user identification. We note that not all encountered users who may be trusted/non-trusted may have been marked. Also only the discoverable bluetooth devices are captured, trusted users that do not have discoverable bluetooth may not have been captured. This issue will be of lesser concern as the adoption of *ConnectEnc* increases.

We rated the performance of *ConnectEnc* for each of the 5 filters (including CF with equal weights) on 2 metrics, 1: number of trusted user in range top 1 to 10, 11 to 20, etc (also known as Precision metric in Information Retrieval literature) and 2. fraction of encounter users needed (from top) to capture 'x'% of trusted users for each filter. The above metrics are chosen to measure how well the filters perform when compared to user's selection. Here ranking is based on the filter score.

For metric 1, we note that *ConnectEnc* is able give high ranks to trusted users (Fig.3A.) . On average, out of top 10 ranked users by FE, DE and CF, 5 (50%) or more users are marked trusted. We see that LV filter's top 10 ranks have 3 to 4 users on average, however, if we consider top 20 users, all filters capture 6-8 trusted users (more than 50% of the total trusted users). The number of trusted user in rest of the ranges continue to fall except in the last range

as it contains all the users ranked beyond 80. For all the filters, there is a strong statistically significant correlation between the score and the rank of trusted users (e.g., for LVC, r=0.84, p <0.01). This shows that a user's willingness to trust others in a mobile network statistically correlates with their behavioral similarity as captured by *ConnectEnc*. Evaluations using metrics 2 shows that 80% of the trusted users are captured by top 25% of the encountering user as ranked by the filters and their is a strong statically significant correlation (Fig. 3B.). Two more metrics are discussed here [15]. We also note that there are users who have high rank, yet they are not trusted. We believe, these can be the encountered users, who are very similar to the user and can provide new interaction opportunities to the user.

Other observations from the deployment include that almost 70% user preferred using equal weights for the *CF* filter. The amount of storage used by the application, on average was 6.2 MB, with storage of filter scores taking only 98KB, rest was occupied by the encounter traces. This shows that storage overhead of *ConnectEnc* filters is quite small when compared to the raw traces. The raw traces can be removed from the device after processing to save space. Also at this rate, 75 MB is needed for storing traces for the whole year.

**Energy Efficiency** Scanning of Bluetooth and WiFi devices consumes maximum power (since the scanning process is periodic). After receiving the traces (which were scanned at 1 min interval), we noted that due to spatial locality in the traces, we can skip the scanning rounds if we find the same devices again in the next round, assuming that the user is in same location with same devices. The number of rounds we skip is $(2^n - 1)$, where $n$ is number of times same devices are found consecutively, with an upper threshold (MaxThres). If, on the next scan round, the devices change, we make $n = 0$ and start again. We note that reducing scanning period increases the loss of encounter information. Since we have the ground truth (traces scanned at 1 min), we can find out the information loss using L1 norm on the distribution of AP (Wifi trace) and Bluetooth devices for both the cases. We note that $n = 2$, gives us 64% saving in scanning, yet the loss is of 6.5%(more in Tab. 1). At n=3 we lose 10% of information but save over 75% of scanning power. We plan to implement this energy efficient scanning algorithm in the future release of *ConnectEnc*, with user-configurable value of $n$.

| MaxThres | Loss(W)% | Saving(W)% | Loss(B)% | Saving(B)% |
|---|---|---|---|---|
| 3 | 6.52 | 64.21 | 6.79 | 66.31 |
| 7 | 10.52 | 75.27 | 11.40 | 76.61 |
| 15 | 15.11 | 81.53 | 15.02 | 82.29 |

**Table 1: Tradeoff between saving in terms of scans and loss of information, W and B indicates Wifi and Bluetooth trace resp.**

# 6. CONCLUSION AND FUTURE WORK

This work introduces, *ConnectEnc*, an effective encounter based framework for establishment of trust in mobile communities in an efficient, privacy-preserving and resilient manner. *ConnectEnc* is driven by trust adviser filters that leverage increased sensing capabilities of the mobile devices and their close association with users, which enables them to capture behavioral similarity with encountered devices and assess trustworthiness levels.

We use four novel trust adviser filters, based on encounter frequency, duration, location behavior-vector and behavior-matrix to generate trust recommendations. The score reflects the level of similarity to aid the user to choose trustworthy nodes in coordination with personal preferences, location priorities, contextual information and/or encounter based keys. The calculations are fully distributed eliminating the need for any server or trusted third party. The participatory experiments shows that statistically strong correlation exists between the filter scores and the selection of trusted users.

*ConnectEnc* has been designed to inspire several potential applications that can be enabled in future. However, there are a few avenues that require further research. In the future, we plan to address some of these questions such as handling multiple devices belonging to a user or MAC address spoofing (several techniques exist [16]). Future work will include analysis of other filters for measuring behavioral similarities. We also aim to develop and deploy *ConnectEnc* for popular mobile platforms and study the effect of its usage on a larger scale. There is a need to conduct more research in order to understand how trust can be established in mobile societies. We hope that this research contributes to that effort.

# 7. REFERENCES


[1] W. Hsu, D. Dutta, and A. Helmy, "Profile-Cast: Behavior-aware mobile networking," in *IEEE WCNC*, 2008.

[2] M. Mcpherson, L. S. Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[3] C.-H. O. Chen and et al, "Gangs: gather, authenticate 'n group securely," in *MobiCom '08*, 2008.

[4] U. Kumar, G. Thakur, and A. Helmy, "Protect: proximity-based trust-advisor using encounters for mobile societies." in *IWCMC'10*.

[5] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell, "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application," in *SenSys '08*.

[6] N. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. Campbell, "A survey of mobile phone sensing," *Communications Magazine, IEEE*, sept. 2010.

[7] K. K. Rachuri and et al, "Emotionsense: a mobile phones based adaptive platform for experimental social psychology research," in *Ubicomp*, 2010.

[8] "Participatory Sensing," 2012. [Online]. Available: http://participatorysensing.org/

[9] Y.-H. Lin and et al, "Spate: small-group pki-less authenticated trust establishment," in *MobiSys*, 2009.

[10] J. M. McCune, A. Perrig, and M. K. Reiter, "Seeing Is Believing: using camera phones for human authentication," *Int. J. Secur. Netw.*, vol. 4, no. 1/2, pp. 43–56, 2009.

[11] S. Buchegger and J.-Y. Le Boudec, "Self-Policing Mobile Ad-Hoc Networks by Reputation," *IEEE Comm. Mag.*, vol. 43, no. 7, p. 101, 2005.

[12] S. Buchegger and J.-Y. L. Boudec, "A robust reputation system for mobile ad-hoc networks," in *P2PEcon*, 2003.

[13] S. Zhong, J. Chen, and R. Yang, "Sprite: A Simple, Cheat-Proof, Credit-Based System for Mobile Ad-Hoc Networks," in *INFOCOM*, 2002.

[14] R. S. Burt, "Decay functions," *Social Networks*, vol. 22, no. 1, pp. 1 – 28, 2000.

[15] "Supplementary Information." [Online]. Available: http://www.cise.ufl.edu/tr/REP-2012-543/

[16] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks," *Wireless Communications*, 2010.