Privacy-preserving Adsense Systems Using Delay Tolerant Networking

Yixuan Fan
Beijing University of Posts
and Telecommunications

Hamed Haddadi
Queen Mary
University of London

Pan Hui
Telekom Innovation
Laboratories

ABSTRACT

With the ever-increasing number of smart phones, a growing numbers of people view advertisements on their phones and hence the smart phone advertising market has become rich and noticeable. To raise click-through rate and maximize profit, ad brokers ensure their ads are more personalized and targeted. Therefore, they collect personal information to build an accurate user profile. The use of sensitive and personal information may raise privacy concerns. In this paper we focus using Delay Tolerant Networking (DTN) to anonymize click reports, aiming to stop attackers tracking and identifying users based on behaviour and location. The results of our simulations prove that a few-hop DTN-based system can protect users' identity and privacy while not heavily increasing their energy costs.

1. INTRODUCTION

Online adverting has become one of the largest revenue sources of many Internet giants, such as Google, Microsoft and Apple. Google's advertising revenue in 2011 was over $36 billion and is only expected to increase over time.[1] Surveys from Gartner and Telsyte group suggest that nearly a third of mobile phone users are using smart phones. With the ever-increasing number of smart phones, the cell phone advertising market is fertile and noticeable. To promote users click advertisements and maximize their profit, advertising brokers hope their ads are personalized and targeted. Accordingly, they need the vast amount of information kept on phones, like location and interest of the user, to build an accurate user profile and choose appropriate ads to display based on such profile. In order to compensate content provider and bill advertiser, brokers have to collect click reports for feedback after an ad is clicked. The use of sensitive and personal information may raise privacy concerns. Many people have been aware of the fact that user privacy is at risk stemming from the information which is sent to brokers for targeted ads. Here is another fact that user privacy is implied by reports. However, this problem does not get enough attention. For example, the content of click reports may be disclosed to unauthorized parties. Alternatively, user behaviour and location information may be used to trace and infer communication patterns. Some of these privacy concerns may be mitigated by technical methods. For instance, Public key encryption can be employed in preventing unintended parties from reading click reports. And permission requirements can protect users privacy and users can more actively control the information according to clear expectations about how their information is used.

Our paper mainly focuses on how to prevent tracing and inferring identity of user on the basis of user behaviour and location information revealed by click reports. Some current privacy-aware applications neglect the importance of user behaviour and even do not regard user behaviour as privacy. The key point of our project is using Delay-Tolerant Networking (DTN) to anonymize reports before being transmitted to Ad Provider. Although the content of reports can be encrypted and unreadable for others, traditional cellular network may still leak users' privacy. More private information can be inferred in accordance with physical location of the base station where reports are sent. Through long-term monitoring, the places which are frequently visited is high likely to be users' home or work place, and brokers may deduce that a consumer who has some kinds of hobby lives in a community from which reports about certain ads are always sent. So, usually, users do not want their physical location and behaviour is revealed by reports. The purpose of anonymous reports is to blur users' identity and preserve their privacy. In the anonymous progress, DTN replaces traditional network. There are several intermediate relays which can provide enough delay and restrict tracing. In this paper work, we first design a suitable system for anonymous click reports. Then we perform simulation under different conditions. Finally, based on the results, We analyse the feasibility of the system and feasibility circumstance. We focus on two primary parameters, the success rate of transmission and the power consumption for each report. Our main achievements are simulating the system using DTN to anonymize click reports and giving a feasibility analysis of this system. Our results prove that a few-hop DTN-based system can protect users' identity and privacy while not heavily increasing their energy costs.

2. BACKGROUND

2.1 The overview of advertising architecture

In order to better understand the advertising system, We would like to introduce some important roles of the system and their relationship.

- Advertiser: Advertisers provide advertisement texts, targeting information and other relevant data which aim to reach a specific group of users depending on requirements such as gender, age, interests and location and promote these users purchase specific products. In general, advertisers do not show their ads to users directly. They employ ad providers to display ads on websites.
- Content Provider: They are also called publishers. Content providers have their own websites, such as news sites and blogs, whose main purpose is not advertising. They provide online services or products to attract users. Alongside their main content, several advertising boxes are used to display targeted ads which are of interest to users. To some extent, the content of websites may be helpful in targeting. The number of ads displayed and the number of clicks on the ads determine the profit of content providers.

[1] http://investor.google.com/financial/tables.html

- Ad Provider: Ad providers, such as Google, Microsoft and Apple, are also called brokers who are the interface between the advertisers and content providers. They organize whole advertisement network and control advertising process. They gather ads from advertisers, provide ads for the users on the publisher websites, collect click reports, bill the advertisers and pay the publishers on the basis of reports collected. They want to maximize their own profit and keep advertisers happy.
- Users: They view ads displayed on publisher websites and buy the products that suits them. When a user clicks advertisement, a click report will be sent to ad provider. End users simply want to discover interesting products with relatively small cost including battery life and personal information.

Indeed, there are also some other components in advertising network, such as Network Operator, Profiling Agent and so on. We do not introduce them since they are irrelevant to our system.

2.2 Current advertising systems

2.2.1 Privacy-aware systems

Nowadays, detailed profiling and data-mining techniques may result in a viewpoint that targeting is inherently in conflict with privacy. Nevertheless, the concept of mobile or pervasive advertising has been researched for several years. Many researchers argue that other alternatives exist [3].

Adnostic [8] proposes client-side software which profiles the user and keeps the profile secret. When a user views a webpage and requests ads, the broker selects a group of ads (they recommend 20) which fit well with the ad page and sends all of them to the client. Then, the client chooses the most suitable ad from this group to show the user. Adnostic uses homomorphic encryption and efficient zero-knowledge proofs to let the broker add up the number of clicks for each ad without sensing the results which remain encrypted. Instead, a trusted third-party (TTP) can decrypt results. Adnostic does not treat users' browsing behaviour and click behaviour as privacy. In fact, knowledge of which ads a user has clicked on and location demographics allows the broker to identify the user. Privad [1] has a similar purpose as Adnostic but its design is quite different. The novel aspect of Privad is dealer that is a new entity, like an intermediate, and anonymizes the client to prevent the origin of the clicks being easily traced. Communications between the client and the broker are encrypted with the broker's public key. The dealer has so much work, deciding which ad should be displayed, removing the identity of reports, detecting fraud clicks and defending replay attacks. It may become a bottleneck since too much information stored in it.

Although these designs are privacy-aware, they are not designed for mobile advertising and hence do not consider the characteristics of mobile environment.

2.2.2 Payment model

Another significant issue is payment model. Pay-per-click (PPC) is the simplest method that even does not need feedback report. Content sites commonly charge a fixed price per click rather than use a bidding system. PPC implements a model called affiliate and offers financial incentives (in the form of a percentage of revenue) to affiliated partner (content sites). However, the PPC advertising model is open to abuse through fraud click. It is easy for botnet to attack this model. Juels [5] proposed a cryptographic approach to replacing the pay-per-click model with pay-per-action (PPA). *Action*â that may be shopping, login or form submission helps to distinguish unsuccessful clicks which are discarded. In this model, the users who make a purchase are identified and use a coupon instantiated by third party cookies or issued by the attester (the ad broker) on redirection. The weak point of this model is malicious advertiser. The malicious advertiser intentionally use a botnet to replay the coupons numerous times and let the broker detect this replay attack, thereby discounting all these replays or removing these clients from the system with valid coupons. The income of broker may be minimized. Despite the merits of this method, it has not been implemented on a large scale as it requires trust between advertisers and publishers. Two substitutes, cost-per-impression or cost-per-click, are generally employed to bill advertisers. In contrast with PPC and PPA, they need click reports for feedback. The purposes of collecting click reports are to charge the advertiser and pay the content provider. To protect the content of reports, reports are encrypted. To further protect users against attempts to link reports to user behaviour, a new approach, DTN, should be taken.

3. DESIGN AND IMPLEMENTATION

3.1 The design of system

To begin with, We would like to give you a fundamental idea about our system. In order to minimize the possibility of touching information, each mobile phone should have an application that supports the advertising system. It ensures there are only necessary parties in reports transmission process. The user who joins in the DTN agrees to relay others reports to advertising provider, once this application is installed. Such application is responsible to generate and encrypt click report automatically when user views and clicks interesting advertisement.

There are three essential requirements due to privacy consideration [4]:

- The relays should have certain social in-correlation with the social network, which prevents identity reverse engineering from the social relationship.
- If possible, the final location of the final hop to the cellular network should have certain geographical distance from the original location of the report.
- If possible, there should be a certain delay between the time when the report is first sent out from the source and the time when it is finally sent to the destination.

In our system, utilizing DTN is able to meet these requirements. Firstly, the source periodically scans the environment. And the relays are chosen from neighbours randomly, so they may be friends or strangers. Generally, they do not have certain social relationship and reverse engineering cannot work. Secondly, there exists a geographical distance between every two relays. After a relay leaves the location where it receives the report a specific distance, it will transmit this report to next relay. Therefore, when the report reaches the final hop, it has a random distance away from the source. In this way, such distance can prevent inferring the original location of the report. Thirdly, to achieve the delay criteria, the system should define what the required proximity is. That is, two mobile phones have an opportunity to exchange their data when their distance is less than a specific value, for a minimum period of time, which should ensure a suitable delay as well as good performance of the whole system. In a word, thanks to DTN, users' identity, location and other information cannot be easily inferred according to the behaviour of reports, thereby preserving users' privacy. Each transmission needs three intermediate relays so as to anonymize the click report. DTN relies on mobile peer-to-peer store-and-forward and hence there is no additional monetary cost on top of the cellular network cost. Depending on the seminal work on 6-degree of separation by Milgram [6], 3 relays (so in total 4 hops) should be long enough for the report to have enough temporal delay and geographical distance from the source, thus scrambling the social relationship.

3.1.1 Other challenges

There remain several security and privacy challenges to be addressed. We have to take some approaches to main threats, such as interception, fraud click and replay attack. Many attackers monitor all the forwarded packets in the network and want to intercept messages. Public key encryption can solve this problem effectively. The click report is encrypted with ad provider public key when it is sent. Only ad provider can decrypt the report and read the content.

Although fraud clicks and replay attacks may not be a privacy problem, they influence the normal operation of system and We have to address them. One of the main reasons why online advertising platforms keep detailed logs of user clicks on ads is to detect instances of fraud. To detect fraud clicks, Bluff ads [2] are chosen randomly and added to the ads displayed to the user. They are real ads but untargeted. Bluff ads should occupy a small part of the publisher page. Too many bluff ads not only deceases ad providers' and publishers' profit but also leads to poor user experience. Consequently, the Bluff/real ratio must be set in a way that the user' browsing experience and advertising quality perception is not greatly affected. Such ratio can help determine whether a user is benign. There are some different forms of click-fraud attacks. Publisher fraud is the most common case. In this way, a publisher employs a large botnet to perform clicks for it. The purpose of bots hired is to increase the number of clicks, and they cannot distinguish Bluff ads and targeted ads. Hence, the Bluff/real ratio of bots is usually higher than an average user. The high Bluff/real ratio and frequent click are good indications of a bot. Moreover, fraud clicks can be used to attack specific target, such as advertiser. These attacks are aimed at adding extra monetary cost of the advertiser since advertiser has to pay each click. Such attack can be realized by a combination of simple threshold sampling and Bluff/real ratio of the advertisements. If most of the clicks from a user are targeted to a single advertiser, there will be an obvious trend in the Bluff/real ratio. In addition, today, advertisers can recognise spammers and click-fraud users via monitor their incoming traffic. If frequent visits from same IP spend no time on the advertiser website, these users will be listed and advertiser will inform the broker removing them from the billing system.

On the other hand, the transmission needs one-time token to avoid reply attacks. Before a user click and view advertisements, the client need to send a request to the broker for a small number of one-time tokens for later sending and forwarding. These tokens should contain a signature (or a one-time pseudorandom number) and later be validated at the broker. The token is unique during a reasonable period. Nevertheless, the broker should not know who it previously gave the tokens to. Users cannot be identified by the unique signature. When a user clicks an advertisement, the source will send an encrypted report with a token. Each time the report is forwarded, a token is added to the report for billing purpose. Therefore, every report that is transmitted successfully should have four tokens which can avoid duplicate report and duplicate payment effectively. Such tokens are used in sending and forwarding reports. We have to achieve a balance between the tokens used for sending and tokens for forwarding. If users send too many reports, tokens for forwarding is lacking and the transmission will be failed. Conversely, lack of tokens for sending reports may prevent users viewing ads and ad provider billing advertise. In general, the ratio of tokens is 1:3, like following Equation.

3.2 The design of click report

Every click transmittable report should have Ad ID, Publisher ID, Relay ID, the Bluff/real ratio for detection, four one-time tokens for validation. Ad ID is used to identify the advertisement which was clicked. Publisher ID indicates the publisher which display the advertisement. Each relay has its own ID and adds Relay ID to the report and encrypts it each time the report is relayed. Ad Provider still use cost-per-click model to bill Advertisers and pay Publishers since it is a better replacement for pay-per-click. Although fraud clicks may exert a negative influence on this payment architecture, Bluff/real ratio can solve this problem. Our system cannot trace users due to lack of a dealer, so the statistic, such as the Bluff/real ratio, should be done by the client-side and then delivered result to Ad Provider. If the ratio is higher than a threshold, the report should be ignored. Then the broker receives the report and checks the token included in the report. If the report has four valid tokens, the broker will accept it and offer incentives, such as discount of monthly fees, to those relays according to Relay ID.

3.3 The design of simulation program

For our simulations, we use Brownian motion model for users' mobility. This is a naive model, however the focus of paper is not mobility models. Rather, we wish to establish the lower bounds on privacy using click report handover. Use of realistic mobility traces will leverage on individuals contacts at points of interests such as train stations and common route. However, such *realistic* city-wide mobility traces have not been made publicly available and we leave that to future work. The program consists of two components: 1.Generating the route map for users, and 2. Simulating model in the map. The system scans the environment and updates users' position per hour. All users' x and y coordinates are randomly generated and they form the routes. The random movement of users follows the normal distribution. Most displace is less than 2 km which make the system more realistic. After all, the scope of action of most people is relatively small.Before simulating the system in various scenarios, some parameters should be initialized, such as area of region, the number of users and Time-To-Live (TTL). Our computer, exactly the memory, limits the scale of simulation. In fact, the area cannot exceed $100km^2$. The coverage region should be a single region rather than multiple region, which is constructed of DTN hosts and DTN relays but no gateway. In this model, delay is not a main concern for the delivery of the advertisement reports. Taking account of battery life, frequent wireless scan may reduce the efficiency of DTN routing. Therefore, each phone scans the environment and detects neighbours once per hour and has opportunities to contact with each other. In our program, the coverage region (the scope that users randomly move), the total number of users and the forward distance between two relays are variable and will be discussed later. The constant parameters are the TTL, the communication distance between two mobile phones, the number of tokens that each user gets per day and the power consumption per hour. There is a paradox. For getting enough delay, TTL cannot be too short. Nevertheless, too long TTL may hinder billing process. So 7-day is a reasonable and feasible TTL.

Users in this system use Bluetooth to link with each other and exchange data directly when in proximity. The Bluetooth in cell phone usually is Class 2. That is, the communication distance is around 10 meters. Each user acquires 10 tokens for clicking ads from ad provider every day. The user can view and click ad anytime and use a token for sending click report. Lastly, We explain the power consumption per hour. Although many factors influence energy consumption in the real world, We assume it is only related to transmission time. The more time the transmission takes

Scenario	A	B	C
Area (km^2)	10x 10	10x 10	10x 10
Total Number	20,000	30,000	40,000
Density (users/km^2)	200	300	400

Table 1: User Density

the more energy the report consumes. So, in our design, the power consumption can indirectly reflect how long the report transmission takes to reach ad provider from the source user. The battery capacity of smart phone is generally 1000 1500mAh. As the Bluetooth is in holding state, the energy cost is about 1mA.Though the cost of Bluetooth transmitting data (working state) may reach 100mA, the duration (peak time) is too short so the main cost of Bluetooth is from holding state. Accordingly, We set the consumption of Bluetooth to be 0.1% battery per hour. Before forwarding the report, the TTL of report should be checked first. And next relay must have enough tokens and battery available for forward.

Next, we will determine the density range of users in this field. The density influences the probability that users exchange reports and controls the delay. Two requirements need to be met. On one hand, each report should have temporal delay. Reports will be forwarded immediately once there are available relays due to flooding algorithm. So each user should have the limited probability of meeting others at certain time. On the other hand, the number of users cannot be too small. Without enough users, reports cannot be transmitted and the system fails. Therefore, the ideal situation is that the report is forwarded with proper delay. As shown in Table1, We assume three different densities, A, B and C, and the region that the system covers is square with an area of 100 km^2, and the distribution of users is uniform. We draw two pictures about probability of users within the communication range according to Poisson distribution. The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.

The communication distance is $r = 0.01$ km and the communication area is

$$S = \pi r^2 \tag{1}$$

S is about $0.000314 km^2$. The density of users is λ. The probability, $p(n)$, which n users exist in S is

$$p(n) = \frac{(\lambda S)^n}{n!} e^- \lambda S \tag{2}$$

Hence the probability, $q(k)$, which at least k users are in S is

$$q(k) = 1 - \sum_{n=0}^{k_1} \frac{(\lambda S)^n}{n!} e^- \lambda S \tag{3}$$

As shown in Figure 1, within one hour, the probability that there is one user in S is far less than 1. In other words, there may be no relay in the communication region. Naturally, the encounter probability is proportional to the density. If the density of users is larger than 10,000, the probability that one user is in communication region is 100%. That is, the report will be forwarded to next relay without delay.

A report should be received by the broker within 7 days through 3 relays. On average, arriving each relay spends about $\frac{TTL}{3}$. When the density of users is larger than 300 and less than 10,000, the report can be successfully forwarded with certain delay. Lower density needs longer communication distance and so Bluetooth may be not qualified. However, 10,000 users per square kilometre may be too large. Beijing may be a good instance. The urban area of Beijing is about 1,000 km^2. Beijing has a population of more than 10 million. Beijing has more than 10 million mobile phone users and nearly one-third of them use smart phone that can support our system. The density of smart phone users is about 3,000 $users/km^2$. Many of those users are not willing to use the system due to many reasons. For example, some of them may view and click advertisements occasionally. Some of them may think installing an application to preserve privacy is unnecessary or even pay no attention to their personal information.

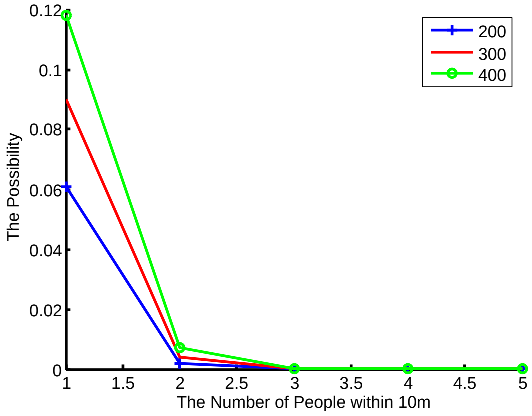

Figure 1: The encounter probabilities within one hour under three conditions.

4. RESULTS AND EVALUATIONS

We concentrate on discussing other three main factors that influence the performance of system, the success rate of report transmission and the power consumption of each report. The higher the success rate is and the less transmission cost, the better the system performs.

1. The area of coverage region (km^2)
2. The user density (users/km^2)
3. The forward distance (km)

In the simulation, the square coverage region where all users move randomly may be 1 to $100km^2$. As described above, the density should be 300 to 500 per km^2. Once the distance from the location where a relay receives the report is larger than the forward distance, the report will be forwarded to next relay. Such distance should be as far as possible so that user can hide identity and prevent tracing, but too large distance may be inappropriate and hinder transmission. For instance, the coverage region is $5x5$km and its diagonal is about 7 km. The forward distance 10 km may be overly large for the region $5x5$. As a result, the forward distance is determined on the basis of the area and their relationship will be discussed later. Because of these different situations, so hundreds or thousands of simulations should be done. Although simulating all conditions can get more reliable results, the workload is so heavy and consuming time so much that it is actually difficult to achieve. In this case, orthogonal experiment is a practicable way which can

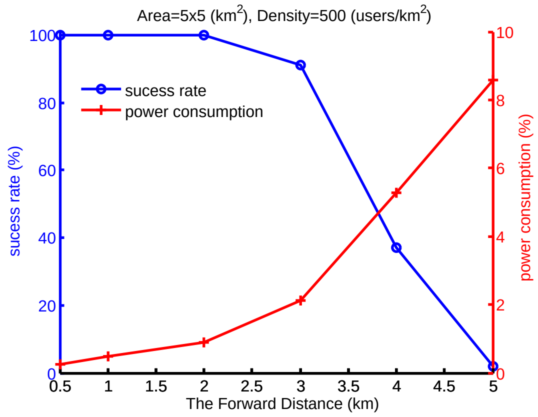

Figure 2: Different performances when the area equals 25 km^2, the user density is 500 users/km^2, the forward distance hovers around half of the side length.

decrease the number of simulations effectively. We can understand all results and find out the optimal combination via analysing several typical results.

4.1 The feasible and optimum condition

4.1.1 The forward distance

Firstly, we test the performance with the change of primary factor, the forward distance. The distance varies from 0.5 km to 5 km. Two other factors are invariable. 12,500 users are in 25 km^2 (5x5). As we known, the optimal forward distance may be different as the area of coverage region changes. There may be a relationship between them. The line of success rate declines obviously at 3 km. At the same time, a sharp grow in power consumption arises. The percentage of battery utilized for sending and receiving report is almost no change. Hence, the main cause of increase of power consumption is longer detection since relays have to spend more time looking for next relay until TTL becomes zero. Based on the trend presented in Table1 and Figure 2, the forward distance which equals about half of the side length of coverage region is a divide. Once the distance exceeds 3 km, the performance drops dramatically. More reports are discarded because of timeout and each report, on average, spends more time detecting available relay and consumes more power for transfer. Accordingly, we guess the distance which is less than half of the side length of field enables reports to be received by the destination with less consumption. In this context, the distance should be as far as possible, thereby confusing tracer and protecting privacy.

We also did two other similar simulations to prove our conjecture. The values of area are 7x7 and 10x10. The user density is still 500. The values of forward distance hover around half of the side length. As we known, the optimal forward distance may be different as the area of coverage region changes. There may be a relationship between them. The line of success rate declines sharply at 3 km. At the same time, a sharp grow in power consumption arises. The percentage of battery utilized for sending and receiving reports experiences nearly no change.

As can be seen from Table 2 and relevant Figures 3 and 4, the forward distance is an influential factor. The forward distance is proportional to the power consumption but inversely proportional to the success rate. The success rate keeps 100% and power consumption is relatively low until the distance is larger than half of the side length. Here the distance that is less than half of the side length is regarded as optimal.

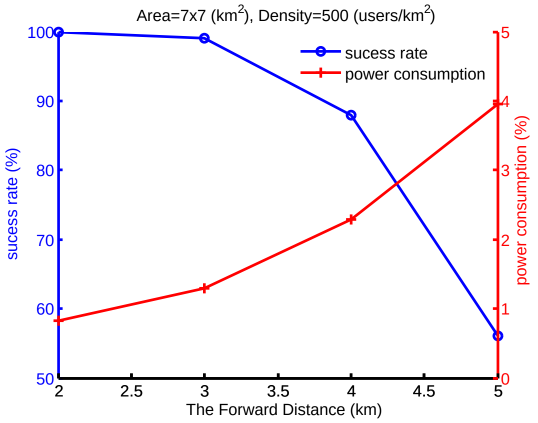

Figure 3: Different performances when the area equals 49 km^2, the user density is 500 uses/km^2, the forward distance hovers around half of the side length.

4.1.2 The coverage region

Then, in order to looking for the optimum of secondary factor-the coverage region, We still keep the user density constant with the change of area. And the forward distance should be the ideal value- less than half of the side length of field. It is clear that the line of success rate fluctuate around 100%. Although it slightly decreases to 99% sometimes, We tend to ignore hundreds failed reports compared to the hundreds of the thousands of reports sent. Besides, there is an interesting wave about power consumption for each report. The wave has increased pretty drastically, but it has three slight drops. We hold that the cause of increase is the increment of forward distance and the cause of drop is the change of area of the coverage region. When two simulations have same forward distance, , larger area results in less consumption. Thus, the impact of coverage region is not noticeable, in particular when the distance is optimal. Here, the 2 km forward distance is a good choice for all possible coverage regions.

4.1.3 The user density

Finally, we study the user density. Its range is relatively small, so We did three similar simulations. In each simulation, the values of coverage region and forward distance are constant and ideal. And the density grows from 300 to 500. As the user density increases, the success rate of transmission gains and power consumption decreases. Hence, the user density should be as large as possible. After all, more people (mobile phones) means more opportunities to contact and exchange data. Here 500 users/km^2 is optimal and this is shown in Figure 5.

The forward distance is proportional to the power consumption but inversely proportional to the success rate. The success rate keeps 100% and power consumption is relatively low until the distance is larger than half of the side length. Here the distance that is less than half of the side length is regarded as optimal.

5. CONCLUSION

Our results reveal a trend that the success rate of transmission and consumption are in inverse proportion. The success rate de-

No.	Area (km^2)	Density ($users/km^2$)	Distance (km)	Success Rate (%)	Power Consumption (%)
1	5 x 5	500	0.5	100	0.25
2	5 x 5	500	1	100	0.47
3	5 x 5	500	2	100	0.89
4	5 x 5	500	3	91	2.13
5	5 x 5	500	4	37	5.28
6	5 x 5	500	5	2	8.59

Table 2: Performances with different forward distance

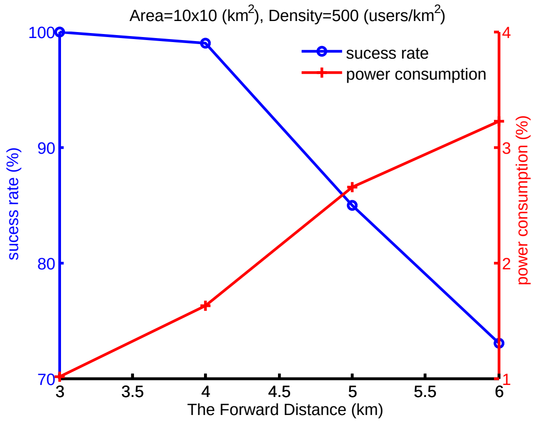

Figure 4: Different performances when the area equals 100 km^2, the user density is 500 uses/km^2, the forward distance hovers around half of the side length.

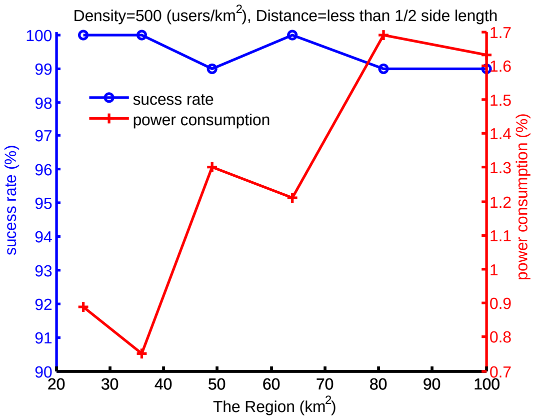

Figure 5: Different performances when the user density is 500 uses/km^2, the forward distance is less than the side length, the area changes from 25 to 100 km^2.

clines while much time and power are demanded. In addition, We would like to explain the impact of the user density, the area of square coverage region and the forward distance on performance. Greater user density makes performance better, the success rate higher and consumption less. The area of coverage region, the forward distance and performance have a more complicated relationship. Taking account of protecting users' privacy, we hope the forward distance is as far as possible. The area of square coverage region, nevertheless, limits the forward distance and poor performance result from too large distance. We need to strike a balance. The forward distance should be less than half of the side length of coverage region and far larger than the communication distance, thereby guaranteeing good performance as well as preventing linking reports to user behaviour. Under appropriate conditions, all reports can be sent to destination via three intermediate relays and some hours delay. These relays and delay can protect user behaviour. The power consumption does not exceed 3% of battery (delay within 30 hours) and is acceptable. In conclusion, this system is feasible and scalable. It improves privacy without increasing monetary cost of the cellular network.

Real world models of user distribution, user movement and the data about battery consumptions are required to improve these results. In reality, users' distribution across a town are never random and daily movement route follow specific patterns. However such patterns would most likely enhance the performance of our prototype. Using mobility models such as those discovered in [7] will improve these results. Battery usage is another issue. In these simulation, we assume the battery consumption is only a factor of time while In reality, many factors lead to change in energy consumption.

6. REFERENCES

[1] S. Guha, A. Reznichenko, K. Tang, H. Haddadi, and P. Francis. Serving ads from localhost for performance, privacy, and profit. In *ACM Workshop on Hot Topics in Networks*, 2009.

[2] H. Haddadi. Fighting online click-fraud using bluff ads. *ACM Computer Communication Review*, 40(2), 2010.

[3] H. Haddadi, S. Guha, and P. Francis. Not all adware is badware: Towards privacy-aware advertising. In *9th IFIP Conference on e-Business, e-Services and e-Society (I3E)*, Nancy, France, 2009.

[4] H. Haddadi, P. Hui, and I. Brown. Mobiad: private and scalable mobile advertising. MobiArch '10, pages 33–38, New York, NY, USA, 2010. ACM.

[5] A. Juels. Targeted advertising ... and privacy too. In *CT-RSA 2001: Proceedings of the 2001 Conference on Topics in Cryptology*, pages 408–424, London, UK, 2001. Springer-Verlag.

[6] S. Milgram. The small world problem. *Psychology Today*, (2):60–67, 1967.

[7] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil. An empirical study of geographic user activity patterns in foursquare. In *Proc. of the 5th Int'l AAAI Conference on Weblogs and Social Media*, pages 570–573, 2011.

[8] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas. Adnostic: Privacy preserving targeted advertising. In *NDSS 2010*, San Diego, California, USA.