Towards Trusted Online Dissemination of Consumer Information*

Mohamed Eltoweissy  
Department of Computer Science  
Virginia Tech  
Falls Church, VA 22043  
eltoweissy@vt.edu

M. Hossain Heydari  
Department of Computer Science  
James Madison University  
Harrisonburg, VA 22807  
heydarmh@jmu.edu

Linda Morales  
Dept. of Computer Science and Information Systems  
Texas A&M University-Commerce  
Commerce, TX 75429

I. Hal Sudborough  
Department of Computer Science  
University of Texas at Dallas  
Richardson, TX 75083

Abstract

Given the growing number of electronic services and their applications in many existing and evolving domains, such as e-tail and e-medicine, much recent attention has been focused on the collection, maintenance, and dissemination of personal information by electronic service providers (ESPs). Often, information is gathered without the active involvement of the consumers, who may be unaware of the nature of the data collected or the purposes for which these data will be used. Techniques such as cookies generally are implemented on the browsers without permission from the targeted consumer. In this paper, we propose a new framework for information dissemination to ESPs using secure communication of consumer-controlled information via trusted information distribution (TID) services. A TID service gathers information based on consumer consent and consumer-controlled profile then disseminates relevant information to targeted groups of ESPs. Using encryption and a group key management scheme, our communications mechanism securely and discriminately disseminates information to dynamic groups of ESPs that subscribe to the TID service.

Keywords: Electronic Services, Privacy, Group Communications, Combinatorial Optimization, Key Management, Internet Privacy, Consumer Consent

I. Introduction

For retailers and service providers who do business on the Internet, referred to in this paper as electronic service providers (ESPs), information is a prized commodity. ESPs create databases to store data they collect and use data mining and other tools to analyze it. The objective of ESPs is to provide an online personalized buying experience and customized products and services for improved customer satisfaction and retention. This results in increased revenue and leads to economic growth.

Information that ESPs routinely collect about potential consumers falls into three categories: personal data, digital behavior, and communications. Consumers’ personal data include information such as name, marital status, residence, e-mail addresses, phone numbers, financial information, and health information. Digital behavior refers to consumers’ activities online. These include the sites they visit, frequency and duration of these visits, and online shopping patterns. Consumer communications include electronic messages, postings to electronic boards, and votes submitted to online polls and surveys [4].

Several direct and indirect ways are used to collect consumer data. Direct ways include cookies, web beacons, and screen scraping. Indirect ways include deductive analysis on collected data. With the emergence of the web as a powerful e-commerce medium, the buying, selling, and trading of individual and aggregated consumer information has become a mainstream...
industry. However, for e-commerce to realize its full potential preserving consumer privacy is of critical importance. Personal information in the wrong hands can cause financial and social problems for the victims. Identity theft has become a major problem in recent years. Another major impediment to the growth of e-commerce is unsolicited advertisements.

Prompted by a deluge of consumer complaints, the Federal Trade Commission (FTC) created a free national “do-not-call” registry where consumers, who do not want to receive telephone sales calls, can register their home and mobile telephone numbers, via DoNotCall.gov or (888) 382-1222 [6]. A crush of consumers rushed to the federal website launched on June 27th, 2003 to opt-out of telemarketing calls. Reports indicate that, on the first day, they were logging at a rate of 1,000 transactions per second, and 370,000 users had registered by noon. Over 48 million phone numbers have been registered so far and this number is growing. Similar opt-out lists exist for emails and home address that can prevent e-marketers and ESPs from sending advertisements through e-mail and pop-up windows.

The “do-not-call” registry may have a huge negative impact on the e-marketing industry and the ESPs that use their services. We believe that although this approach can limit the number of annoying telephone calls, emails, and popup ads, it is too extreme and counter productive for both the consumers and the ESPs. It is anticipated that over $1 Billion a year will be lost in marketing revenues and over two million jobs will vanish as a result [1][2][7].

The problem with the FTC’s opt-out approach is that it provides an all-or-nothing mechanism. We advocate an approach based on the principles of consent and security. While most consumers find unsolicited ads an annoyance, many may spend hours on the Internet looking for the best deal on a product they are considering buying. This is an opportune time for the ESP’s to advertise these particular products to the individual consumer. A consumer may be more willing to share personal information in order to get information about specific products. For example, information about new medication or breakthroughs in treatment would be welcomed by most who suffer from a disease. Once a product is purchased, consumers, in most cases, would not want to receive further advertisements for the product.

We believe that the adoption of a consent-based model for secure information dissemination would enhance e-service revenue by causing consumers to have greater trust and a more positive attitude toward ESPs. This will result in:

- Consumers purchasing more products and services online, and
- ESPs that adopt a consent-based model, achieving significant positive branding.

In this paper, we propose a Consent-Based Information Dissemination framework, (or CBID), for the secure consent-based dissemination of consumer information. The proposed framework enables a consumer to control the information sent to ESPs about them and to elect to receive only the advertisements and services that they are interested in at the time. Each consumer creates a profile that indicates the information and/or services they are interested in at the time and what part of their personal information can be disseminated to which ESP. Consumers can change their profiles to request different information and/or service as they need.

For CBID to succeed, information about consumers must be conveyed to the targeted ESPs both securely and efficiently. Security is important so that consumers can trust the system, while efficiency is important so that information is disseminated to the intended ESPs quickly. This would result in consumers receiving requested advertisements and customized services and products in a timely manner. In addition, the framework must scale well with respect to the number of consumers and the number of ESPs.

Our proposed framework uses Trusted Information Distribution (TID) services to securely multicast consumer information to subscriber ESPs. In this paper, we describe the components of the framework and their interactions, and we present our solution for securely disseminating information to ESPs. Our security solution is based on the encryption of consumer information with cryptographic keys, and the dissemination of the encrypted information using multicast communications between the TID services and ESP groups. The nature of the system, where ESPs can join and leave dynamically, requires managing the encryption keys to guarantee confidentiality, forward secrecy, and backward secrecy. To achieve this objective, we employ a key management solution, called Exclusion Basis Systems (EBS) that we developed in [3]. EBS is a combinatorial formulation of the group key management problem that yields efficient and scalable key management as will be described in Section III.

II. Consent-based Information Dissemination Framework (CBID)

Our system model includes consumers, Trusted Information Distribution (TID) services, and E-Service Providers (ESPs). These entities interact as depicted in
Individual consumers may subscribe to many TID services, and each ESP may also subscribe to many TID services. A consumer creates a profile to describe what kind of data can be collected about him and which parts of their personal information can be shared with the ESPs. TID services are trusted parties that use consumer profiles with consumer consent, and supply the requested information to the subscribing ESPs. ESPs receive consumer information and interact with the consumer regarding the requested information and/or service.

Using secure multicast communications, a TID service sends requested consumer information, approved by the consumer, to groups of the ESPs subscribed to their service. As shown in Figure 1, consumer queries can refer to individual ESPs (for example, a request for information from IBM), groups of ESPs that offer the same products and services (for example, all BMW car dealerships in the area), or ESPs that offer overlapping products and services with other ESPs. (Jiffy Lube and Midas offer oil change but Jiffy Lube doesn't work on mufflers). We now describe the main components of the model and their interactions.

![Figure 1: Information Dissemination Framework](image)

### II.1 Consumers

Each consumer creates a consumer consent profile, which will be sent to the selected TIDs. Information in the consumer profile can include consumer’s:

1. Name
2. Telephone Number
3. Home/Work address
4. List of items for which the consumer is interested in receiving information and/or advertisements
5. List of TIDs the consumer has subscribed to
6. Other information

Each field in the profile has a checkmark, indicating which fields of the profile can be available to the ESPs. Each consumer may subscribe to one or more TIDs. Consumers can select which TIDs can receive this information and can edit their profiles as their needs change.

Consumers securely interact with TIDs, using Secure Socket Layer (SSL) and other techniques, to:

1. Receive subscription/payment notification
2. Update their profile
3. Receive alerts on relevant ESP services and products
4. Give or revoke permission for data collection
5. View/edit ESP requested response log
6. View/modify data and profile log
7. Request ESP information

Consumers interact directly with ESPs to:

1. Receive product and/or service information
2. Request more information from the ESPs
3. Purchase services/goods

This interaction (advertisements, catalogues, telephone calls, emails) takes place directly, between the consumer and ESPs, without the TID involvement. Note that the consumer receives only the information he/she requested from the desired ESPs. This interaction is not depicted in Figure 1.

### II.2 Trusted Information Distribution (TID) services

TID services receive consumer profiles from their subscribing consumers and send the appropriate information from the consumer profiles to the subscribing ESPs. TID services are the trusted party; hence they audit activities of their subscribing ESPs, making sure that the ESPs don’t violate the terms of agreement specified in the consumer profiles. TID services use secure multicast communication to communicate with ESPs. TID services communicate with ESPs to:

1. Send subscription/payment notification
2. Update ESP information
3. Request information on behalf of the individual consumers
4. Receive ESP profiles and data
5. Audit ESP consumer-related information flow

A TID service may be implemented either as a 3rd party service or as a service residing on a consumer’s machine. An analysis of the merits of each of these approaches is a related area of research.

### II.3 Electronic Service Providers (ESPs)
ESP s are companies and organizations that offer products and/or services to the consumers, based on consumer requests. Products and services include advertisements, catalogues, price quotes, product information, and goods and other services. ESPs subscribe to the TID services and once registered they agree to use consumer information only in the way specified by the consumers. TIDs services audit their subscribed ESPs at regular intervals to make sure they do not violate the consumer trust placed on the TID services.

III. Secure Dissemination of Consumer Information

Besides consent-based dissemination, in order for the consumer to trust the TID service, information must be disseminated in a secure way. The security solution in this paper uses cryptographic keys to encrypt communications, and an Exclusion Basis System (EBS), which uses a combinatorial formulation of the group key management problem, to manage the encryption keys. EBS produces optimal results with respect to the parameters \( n, k \) and \( m \), where \( n \) is the size of the group, \( k \) is the number of keys stored by each group member, and \( m \) is the number of re-key messages.

In this Section, we present an overview of EBS. For more details, the reader is referred to [3].

III.1 Overview of Exclusion Basis Systems

An EBS is defined as a collection \( \Gamma \) of subsets of the set of members. Each subset corresponds to a key and the elements of a subset \( A \in \Gamma \) are the members that have that key. An EBS \( \Gamma \) of dimension \( (n,k,m) \) represents a situation in a secure group where there are \( n \) members numbered 1 through \( n \), and where a key server holds a distinct key for each subset in \( \Gamma \). In this section, we will use the terms “key” and “subset” interchangeably. If the subset \( A_i \) is in \( \Gamma \), then the key \( A_i \) is known by each of the members whose number appears in the subset \( A_i \). Furthermore, for each \( t \in \{1, n\} \) there are \( m \) elements in \( \Gamma \) whose union is \( \{1, n\} \) \( \setminus \{t\} \). From this it follows that the key server can evict any member \( t \), re-key, and let all remaining members know the replacement keys for the \( k \) keys they are entitled to know, by multicasting \( m \) messages encrypted by the keys corresponding to the \( m \) elements in \( \Gamma \) whose union is \( \{1, n\} \) \( \setminus \{t\} \). Each new key is encrypted by its predecessor to limit decipherability only to the appropriate members.

To construct \( EBS(n,k,m) \) for feasible \( n, k \) and \( m \), we employ a canonical enumeration of all possible ways of forming subsets of \( k \) objects from a set of \( k + m \) objects. We do this because our construction is based on such an enumeration. There are several algorithms for producing such a sequential enumeration. We choose an enumeration where each element of the sequence is a bit string of length \( k + m \), where a 1 in the \( i^{th} \) position of a string means that object \( i \) is included in that subset, for all \( i \) \( (1 \leq i \leq k + m) \). Note that, every bit string in this enumeration will have exactly \( k \) ones.

<table>
<thead>
<tr>
<th>M_0</th>
<th>M_1</th>
<th>M_2</th>
<th>M_3</th>
<th>M_4</th>
<th>M_5</th>
<th>M_6</th>
<th>M_7</th>
<th>M_8</th>
<th>M_9</th>
</tr>
</thead>
<tbody>
<tr>
<td>K_0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K_1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K_2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K_3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K_4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>X</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Y</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Z</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 2: The canonical matrix \( A \) for \( EBS(10, 3, 2) \) with session keys \( X \), \( Y \) and \( Z \) assigned to sets of members \( (M_0 - M_9) \)

For any \( k \) and \( m \), let \( \text{Canonical}(k,m) \) be the canonical enumeration of all \( C(k + m, k) \) ways to form a subset of \( k \) elements from a set of \( k + m \) objects. For the sequence of bit strings in \( \text{Canonical}(k,m) \) we form a matrix \( A \), where \( k \) and \( m \) are understood, and whose \( C(k + m, k) \) columns are the successive bit strings of \( k + m \) length, each with \( k \) ones. \( A \) is called the canonical matrix for \( EBS(n,k,m) \). For example, the canonical matrix \( A \) for \( EBS(10, 3, 2) \) contains the enumeration of all \( C(5,3) \) ways to form a subset of 3 keys from 5 keys, \( A \) is shown in the non-shaded (upper) portion of Figure 2.

It has been shown in [3] that an optimum EBS incurs half of the overhead of a binary key tree in terms...
of the number of keys per member and the number of re-key messages.

The basic idea behind the solution we propose in this paper is the use of a EBS(n,k,m) to assign and manage keys. We assume that, in addition to the administrative keys corresponding to subsets in \(\Gamma\), a key server also has one or more session keys known to sets of group members. All group members aware of a particular session key constitute a secure communications group. Members in a secure communications group use the group’s session key to encrypt messages exchanged among group members. A primary advantage of the EBS is that it separates administrative keys from session keys, thus simplifying key management.

A canonical matrix is constructed and used as the basis for key assignment and management. The canonical matrix is augmented with session keys. The shaded area in Figure 2 illustrates three rows of session key assignments; keys \(X, Y\) and \(Z\). From Figure 2, evicting \(M_i\) will result in two re-key messages using \(K_3\) and \(K_5\) to communicate new keys \(K'_1, K'_2, K'_3, K'_4, X', Y', Z'\). Note that, even though member \(M_8\) knows \(K_4\) and \(K_5\), it will not be able to decipher new keys \(X'\) and \(Z'\) because these are encrypted with the old keys \(X'\) and \(Z'\), respectively, which are not known to \(M_8\).

### III.2 Application of Exclusion Basis System to Consent Based Information Dissemination Framework

In CBID, a TID service acts as a key server that is responsible for administrative and session key generation and for management of member admission and eviction. It also manages sub-grouping of ESPs as will be explained shortly.

Upon subscription to a CBID, an ESP is assigned a unique ESP key. This key is used for the initial transmission of administrative and session keys from the TID service to the ESP. The administrative keys are assigned to the new ESP based on an EBS canonical matrix like the one in Figure 2. The matrix is stored at the TID service.

ESP's may be grouped together based on parameters, such as service type, paid subscription fee, AddESP

\[\begin{align*}
(1) & \quad n \leftarrow n+1; \\
(2) & \quad \text{if } i > \left(\begin{array}{c} r \\ q \end{array}\right) \text{ then execute either } \text{Extend0} \text{ or } \text{Extend1}; */\text{Extend0 and Extend1 are used to extend an EBS and are beyond the scope of this paper. Please refer to [3] for a detailed description. */} \\
(3) & \quad \text{let } y \text{ be the binary string consisting of the } i^{\text{th}} \text{ element of } S; \\
(4) & \quad \text{let } z \text{ be the binary string } y \text{ with } x \text{ appended}; \\
(5) & \quad (\text{let } y \text{ consist of the individual bits } y_1 y_2 \cdots y_p, \text{ for some } p); \\
(6) & \quad \text{for each } j \text{ (such that } 1 \leq j \leq p \text{ and the bit } y_j \text{ is } 1) \text{ do } \\
(7) & \quad \text{let } k_j' \text{ be a new key to replace the old key } k_j; \\
(8) & \quad \text{Construct a message } T \text{ which includes key } k_j'; \\
(9) & \quad \text{Encrypt } T \text{ by encrypting } T \text{ with the old key } k_j; \\
(10) & \quad \text{for each } j \text{ (such that } 1 \leq j \leq p \text{ and bit } z_j \text{ is } 1) \text{ do } \\
(11) & \quad \text{Construct a message } U' \text{ by encrypting } U \text{ with the new users personal key; } \\
(12) & \quad i \leftarrow i+1 \\
\end{align*}\]

**Figure 3. A program to add an ESP to an EBS**

calendar interest, and location. Each ESP subgroup is assigned a session key. This session key is used to encrypt the consumer data sent to the ESP subgroup. An ESP may belong to more than one ESP subgroup simultaneously and will hold one session key for each subgroup it belongs to, in addition to \(k\) administrative keys, where \(k\) is the parameter in \(EBS(n,k,m)\).

To ensure backward secrecy, when a new ESP joins a subgroup, the session key(s) for the subgroup(s) in which the new ESP will participate should be changed. For each sub-group that the new ESP joins, the new session key is broadcast to all members of the subgroup encrypted with the old session key. In addition, the administrative keys to be shared by the ESPs need to be changed. The new administrative keys are broadcast encrypted with the old administrative keys and sent to all members of the group using the keys that will not be shared by the new member according to the EBS key assignment. Finally, the new administrative and session keys are transmitted to the new member of the subgroup using the new member’s unique key. A program to add a new ESP is shown in Figure 3.
The program *AddUser*, explained in detail in [3], computes appropriate messages to multicast for rekeying when a user is added. The sequence $S$, used in *AddUser*, is the canonical enumeration of binary strings of length $r$ representing all distinct $r \choose q$ ways of choosing $q$ out of $r$ positions, with the chosen positions indicated by 1 and others by 0. The variable $i$ gives the index of the current element of $S$. When $i$ is not larger than $r \choose q$, there are unused strings in the sequence $S$ that can be used for assigning a unique combination of keys to new users. However, if $i$ is larger than $r \choose q$, then all strings in $S$ have been assigned to current users, and the capacity of the current EBS system has been exhausted. To accommodate more users, the length of the strings must be increased by one. For extending an EBS, the reader is referred to [3].

To ensure forward secrecy, when an ESP leaves or is evicted from a subgroup, all keys known to that ESP must be changed. This takes place using the administrative keys not known to that ESP according to the EBS canonical matrix. When transmitting the new keys to the remaining ESPs, the TID encrypts the new keys with their old counterparts to ensure confidentiality. A program to delete the $i^{th}$ ESP is shown in Figure 4.

**DeleteESP(i)**

1. Let $y_1 \ y_2 \ldots \ y_p$, for some $p$, be the bit sequence corresponding to ESP $i$;
2. For each $j$ ($1 \leq j \leq p$ and such that $y_j = 1$) do
   - Create a new key $k'_j$ to replace the old key $k_j$;
3. For each $m$ ($1 \leq m \leq p$ and such that $y_m = 0$) do
   - create a message $Q$ containing each new key $k'_j$ encrypted by its corresponding old key $k_j$;
4. multicast $Q$ encrypted by key $k_m$

**Figure 4. A program to delete the $i^{th}$ ESP**

**IV. Conclusions**

In order to provide customized services and products, ESPs need to collect information about potential consumers. Information gathering is often accomplished without the active involvement of the consumers, who may be unaware of the nature of the data collected or the purposes for which these data will be used. The current solution of opting-out of telemarketing and e-commerce lists is extreme; it hurts both the businesses and the consumers. We conjecture that consumers may be willing to reveal personal information to trusted entities (TIDs) that then securely disseminate the information (with their consent) to specific ESPs that are audited by the TIDs. In return, consumers will receive the advertisements for products and services they are interested in, only when they are looking for these specific goods and services.

In this paper, we have proposed consent-Based Information Dissemination Framework (CBID) a framework that allows the consumers to securely and judiciously share portions of their personal information with ESPs, and request and receive specific information about the products and services they are interested in, in a timely fashion. Consumers can view and modify their consent profiles, as they desire. This would encourage consumers to share some of their personal information with ESPs, hence stimulating e-commerce and at the same time provide consumers with the latest product information at a desirable cost.

Future work includes refining the framework based on a prototype implementation and use by targeted groups of consumers and ESPs. In addition, the functionality of CBID will be expanded to include a Reputation Service (RS) to make sure that TIDs and ESPs comply with rules and regulations, including protecting consumer’s information, ESP information, TID input, and consumer transactions. Another area of investigation is consumer mobility. Finally, compliance with regulations such as the Health Insurance Portability and Accountability Act (HIPAA) needs to be an integral part of CBID.

**VI. References**