Understanding the Market for RFID Traces

Steffan Baron (sbaron@open-source-consultants.de)
*

Matthias Bauer (matthiasb@acm.org)
†

Bettina Berendt (berendt@wiwi.hu-berlin.de)
Benjamin Fabian (bfabian@wiwi.hu-berlin.de)
Matthias Fischmann (fis@wiwi.hu-berlin.de)
Seda Gürses (seda@wiwi.hu-berlin.de)
‡

Sushmita Swaminathan (sswaminathan@diw.de)
§

1 Introduction

RFID tags are widely held to become ubiquitous in logistics in the foreseeable future [11]. Item-level tagging will also pave the way for ubiquitous computing like in so-called smart homes [13]. Applications often envisioned are intelligent cup-boards or smart fridges that know their contents and order missing goods automatically. In this paper we consider the value of information accumulated by observing tag traces, i.e. information gathered by observing tags over time and locations in this hypothetical future, and the costs of harvesting this information.

Tags are small and extremely cheap passive radio devices that, when probed by a reader, broadcast a globally unique serial number within a short range.[1] In the following, we will only consider tags that can do nothing much beyond that; in particular, a tag cannot keep track of who is reading it and when it is read, or selectively deny reading, or disable itself permanently.[2]

The main and currently most influential industry standards for tags, readers and number schemes are set by EPCglobal [6], a joint venture between EAN/GS1 and the Uniform Code Council (UCC). The *Electronic Product Code (EPC)* is expected to replace the conventional bar code on individual items. For convenience, we use the terminology of EPCglobal. However, the relevance of our arguments only

*Open Source Consultants, Berlin
†Shoestring Foundation, Erlangen
‡Institute of Information Systems, Humboldt-University, Berlin
§Deutsches Institut für Wirtschaftsforschung (DIW)

[1] Vendors are aiming at costs of less than 0.05$ per tag for the mass market. Reading ranges depend on frequency, antenna, surroundings and energy. A conservative estimate of the average effective reading range of the tags considered would be around 2–8m.

[2] Standards for more "concious" tags have been proposed that can block read-outs or even communicate with trusted readers only over encrypted radio links. We will briefly discuss those in a moment.

depends on the ubiquitous presence of passive radio tags that transmit *some* sort of globally unique serial number, and EPC is only one possible implementation.

According to EPCglobal, tags carrying unique EPCs will be attached to almost every object in everyday life. They can even be integrated into clothing or other items, and could be nearly invisible to the human eye. The envisioned ubiquity of these tags raises a lot of privacy issues, see for example [2][10][27]. Tag readers are publicly available and cheap, anyone can read all the tags located within a few meters distance, and with minimal effort maintain traces of the area around her readers (i.e., triples of EPC, reading time, and location of the reader). Where EPCs can be associated with their carriers, there is a massive potential for leakage of personal information. This is at least the case for the retailers' billing systems, or virtually for anybody by means of correlation analysis if, for example, remotely readable electronic identity cards become more common.

Furthermore, to support global information exchange about tagged items, EPC-global plans to roll out a huge distributed database, the EPC Network. Once established for inter-supply chain use in logistics, we expect this project might turn into a global information retrieval network for after-sale services and ubiquitous computing environments using RFID tags for object identification (and implicitly for user identification). Obviously this raises further privacy concerns.

Several countermeasures against this kind of leakage have been proposed, see for example [10][16][24][15][6][12]. They usually consist in enhancements to the tag capabilities that allow for blocking or controlling read-outs. These solutions all are bound to fail to solve the fundamental problem for various reasons: (1) security will be imperfect, leaving room for new and creative cyber attacks; (2) some goods are too cheap for protected tags, and unprotected tags will be used for those; and (3) trust in a reader may be unjustified, either because the reader has been hacked, or because the party who owns it is secretly malicious. Therefore, in the following we are only considering unprotected tags as described above.

A more exotic option is to broadcast privacy policies from a PDA [8]. This PDA would then ask all readers in scanning distance not to notice certain sensitive EPCs, and law enforcement will provide the incentives for reader operators to obey the policies that are expressed.

Any approach to security requires the user to configure a privacy policy, but this has already proven not to work on the web for various reasons[3]. First of all, writing privacy policies is complex and cumbersome. Second, content providers on the web have the power to make it maximally inconvenient for users to choose a more restrictive privacy policy. Usually they do this by setting up their own policy which they refuse to negotiate, leaving it to the user to walk away without access to a site, or change her own preferences. In the envisioned future, the forces architecting EPC-based services will be able to repeat this practice, but then instead of just web browsing, most of the real world as humans experience it will be affected.

Most importantly, the incentives to harvest data are strong for those in the position to do so. At the same time, EPCs can be read without being noticed by anyone, which makes privacy policy enforcement practically impossible. Even if the owner of the tagged good would be notified, there could hardly ever be a proof that a law has been broken until the data is out in the open, in the best case ending in a "promise-to-forget-the-data" order and a small fine.[4] On the other hand, the traces

[3]For examples of the vivid discussion on flaws in approaches like P3P see `http://www.kcoyle.net/p3p.html` and `http://www.epic.org/reports/prettypoorprivacy.html`

[4]The proposed use of DRM systems and Trusted Computing in RFID readers for privacy-

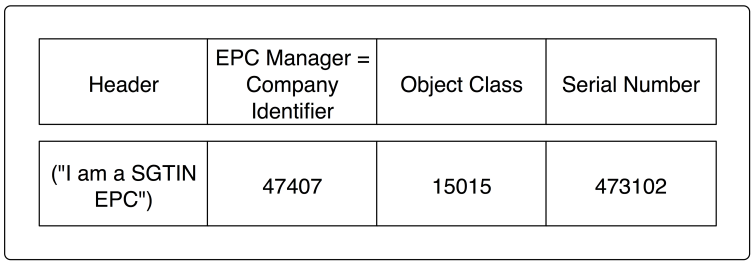

Figure 1: Example Electronic Product Code (EPC)

will contain an abundance of valuable information for many potential buyers. So not only will people produce EPC traces that can be collected unnoticed, but other people will have a strong interest in harvesting and trading these traces.

The aim of this paper is to prove that the incentives in a world with ubiquitous tagging will breed profitable markets for EPC traces, not between the individuals that create them and the corporate world that purchases them, but inside the corporate world, after data on individuals has been harvested.

In Sections 2 and 3, we assess the supply and the demand side, respectively. In Section 4, we discuss the option of changing the incentive structure of the market by tightening privacy laws. In Section 5, we discuss some of the economic and social implications of the EPC trace markets.

2 The Supply Side

The structure of the supply side of a market decides to what extent anybody will engage in production. In our case, the product is a database of the form (i, t, l): i is an identifier (e.g. an arbitrary session ID linking different clicks in a web browsing session, or an EPC); l is a location identifier (a GPS coordinate tuple, an IP address, or a complete URL); and t is a time value (relative or absolute). In this section, we will outline the EPC technology and compare it to the technology of *Web bugs*, that is widely used for online consumer profiling.

2.1 Harvesting

The tags that raise most privacy concerns are those that carry globally unique identifiers, the *Electronic Product Codes (EPC)*. These identifiers are expected to replace the barcode and will serve to uniquely identify any object that carries a tag. The most prevalent kind of EPC has a length of 96 bit and corresponds to a SGTIN (Serialized Global Trade Identification Number). The main structure of this SGTIN-EPC is as follows [6, p.12] (see Figure 1).

The header defines the numeric code in use (here SGTIN); the EPC manager determines which company issued this EPC (usually the manufacturer of the corresponding item); the object class determines the category that the tagged object belongs to, and the serial number identifies a particular item within the same object class.

protection [19] seems attractive but in our opinion suffers from massive scalability and acceptance issues on the infrastructure side.

Note that the serial number enriches the information carried by a tag significantly with respect to a bar code, which only determines the company and object class.

2.1.1 Local EPC Traces

EPC traces (i, t, l) are composed of an EPC i, an absolute time value t, and a location l in any arbitrary but fixed topology (GPS coordinates, internal number of the site of the supermarket, street number and zip code, or the like). The original source of traces is the direct tag-to-reader interaction, where a reader supplies power to all tags in range and collects their EPCs. A reader could be a fixed installation, a hand-held device, or it could be integrated into a mobile phone.

This reader infrastructure could be installed and used officially and with implicit consent, for example inside a shop scanning and tracking the shop's EPCs and those that the customers brought with them as well. It could also be installed for better customer services like style counseling. In ubiquitous computing environments like smart homes, RFID will be a fundamental technology to discriminate contexts, and applications like smart fridges, washing machines and home medical advisors could use it for the inhabitant's benefits. The central privacy problem here is the possible secondary use of these collected traces, which are quite possibly already associated to an individual in the case of smart homes, or very easily linkable in case of shop customers.

There could be rogue readers, mobile or fixed installations, which scan strategic points within cities or buildings to gather traces without consent or notice.

More secretively, there is the possibility to place passive sniffers near official reading infrastructures to capture RFID communications and collect transmitted EPCs.[5]

Finally, creative marketing projects are bound to come up with gadgets like a portable music player that traces tags around its owner all the time, producing a more or less complete profile of her, and even profiles of persons she knows. These mobile trace databases could be uploaded to the vendor and traded for music or other services with insignificant unit production cost.

2.1.2 Harvesting EPC Network Logs

The EPCglobal network is designed to form a global information retrieval network for objects carrying tags with EPCs. It follows the *data-on-network* paradigm, i.e., information about objects is not stored in the corresponding tags. Instead, these tags only contain primary keys for data look-up mostly via Web Services from distributed databases on the Internet, so called EPC Information Services (EPCIS). The main function of the EPC Network is data exchange, first within supply chains, then most probably in the long run as some kind of backbone for an *Internet of Things* and ubiquitous computing environments. In particular, it constitutes an excellent trace trading system.

Every EPCIS can itself be a trace collection point, simply by log file analysis. Instead of the physical location it collects the source IP the query for additional information originated from, and probably even more identifiers and authentication tokens from the information protocol layer, e.g. SQL commands or access control for Web Services. Third parties cannot sniff this trace-containing traffic easily if it is encrypted.

[5] RFIDdump Project: http://rfiddump.org/

But the look-up service *Object Name Service (ONS)* is a weak spot in this design [7]. ONS uses the structure of an EPC, which is transformed into a domain name, for query delegation and is in principle nothing more than classical DNS. So every ONS server in the resolution chain is a potential trace collection point.[6]

ONS is about to inherit all the well-known security weaknesses of DNS, and is, most importantly, a clear text protocol. DNSSEC could be used to mitigate some of these risks, but is until today not globally deployed, and to hope for its widespread use for ONS sometime in the future might turn out as wishful thinking. But even the DNSSEC is explicitly not designed to encrypt communication. So in the worst case, every IP node in the ONS resolution chain, every router, intrusion detection system or network analysis device in the ONS resolution chain might collect traces.

Therefore we argue that the EPC network will make traces at least as abundant as click traces, in addition to local and direct skimming via RFID. Collection and data aggregation systems like Intrusion Detection Systems (IDS) that are in place today could be easily modified to harvest EPC trace data from the Internet. Combining these into large, global aggregation services that exist for e.g. distributed intrusion detection today, or including them in search engines like Google might even make lookups for tags succeed with a high probability.

2.2 Trace Products

We argue that EPC tracing can be turned into a business model similar to those already established for Web click tracing. Web bugs allow site operators to estimate how many users actually viewed an ad and therefore constitutes a popular basis for pricing. This has been very successful for market research purposes, although it has come under heavy attack due to privacy concerns.

In this section, some potential analysis types for EPC tracing will be sketched in order to assess what types of pseudonymity / user identification they require, and whether this is a major difference to Web click traces.

One type of analysis would be analogous to traditional market-basket analysis: finding patterns of co-occurrence, for example between different products or groups of products (each identified by groups of EPCs). This kind of pattern could only be detected through the analysis of mass data; data do not have to be associated with persons, but can be associated with a pseudonym (constructed as a set of locations in close physical proximity, of total size not larger than a person or a person's space such as a car, and moving together).

A type of analysis that is not possible on the Web is a usage analysis that asks to what places (and, if semantic enrichment is available, in what contexts) a particular item has been taken. From this, a profile of a product could be constructed ("this dress is worn at premieres and at the Academy Award Ceremony"), again, pseudonymous association is sufficient, as multiple histories are needed to detect patterns.

To use these patterns for personalization, a carrier of things with EPCs must be identified either as a person, or as a pseudonym in the above sense. The same holds if the purpose is to create the movement profile of a person.

[6]The ONS specification as of today does not use the serial number for delegation, but leaves this option viable in future. But identifying and tracing rare item classes or clusters of items is still easily possible.

All types of analysis appear relevant to market research and customer relationship management in a straightforward way; apart from data integration, the analysis poses no special requirements on processing equipment, thus no prohibitive costs are to be expected.

2.3 Web Bugs in Virtual vs. EPC Traces in Physical Spaces

Navigation in physical space and EPC tracing differe from navigation in virtual space and Web click tracing in several ways.

Data. EPC traces do not show a sequence of volitional events (such that, e.g., a preference for two products must be identified from their relative sequential positions) but a simultaneity of relations to things (e.g., carrying around two products together during all mornings, but never together during the afternoons).

Aggregation Technology. This is a challenge for algorithms: Needed are (a) algorithms on spatio-temporal data of high resolution; (b) data integration to obtain a semantic and thus actionable results; and (c) algorithms that work on data with a rich relational structure.

Such algorithms have been developed over the past (at least) 10 years; they align different granularities of spatial data (e.g., geographical coordinates and street networks), they can integrate the semantics of space and also of action models (e.g., a person's home and work location, different kinds of trips); for a current example, see [18]. There have also been many advances in (multi-)relational data mining, cf. [5]. Thus, if the supplier of EPC traces offers analysis options, the costs for additional intelligence are low. If she only offers data, this cost item does not enter the supplier's cost calculation at all.

The Passive Role of the User. Data come from ubiquitous contexts and not from the interaction with software via an interface that can be configured by the user. By this, we mean the assumption, explained in Section 1, that tags are passive and broadcast in response to requests from all appropriate readers. The consequences are, first, more data, and in principle a higher density of data. On the other hand, the interpretation of what an ID-time-location tripel means may become even harder than it is with click traces, where it can generally not be told why something was chosen. Interpretation is harder because that triple was not a volitional event (choosing to view some content), but the by-product of some other action in physical space. Again, the ensuing additional cost is on smarter data mining algorithms, and the same argument as above holds.

Total Coverage is Harder for Collector. The spatial entities that are being tracked are different, and so is the locus of control. In essence, click tracing starts from *locations* that are comprised of *objects* and that are owned by customers potentially interested in data on these objects (locations are Web sites, objects are products and services offered on these sites, and customers are the site owners). EPC tracing also starts from locations, but these exist independently of the objects that are being tracked, and they are usually public. In the absence of universal tracking coverage, the collection of data that might interest a given customer cannot be guaranteed. This implies additional market risk.

The following approaches appear possible to address the matching problem and integrate the market risk into calculation:

1. The trace collector is the owner of the space under surveillance (see Section 4). In this case, the situation is very similar to the Web situation, the market risk is low.
2. The trace collector may already have entered into a contract with a customer, then search at locations where data related to that customer are likely to accrue. The risk that no interesting data are collected must be borne by either or both of the contract partners, and appropriate pricing models be developed. To be able to identify potential customers, the trace collector must have access to data sources that resolve EPCs to manufacturers or retailers, e.g., the EPC network. If this is not possible, then the trace collector can only offer rich search functionality (what IDs have been tracked) and hope that customers search for these data and express their interest in buying.

Thus, EPC tracing is likely to have a higher market risk than Web click tracing. Quantifying this risk is a task for future research.

Summing up, EPC tracing appears to imply few additional costs relative to Web click tracing. One of the major cost items may be the market risk. It is difficult to quantify this market risk, but as it hinges critically on the awareness of EPC tracing in general and of the specific trace collector in particular, a small number of successful applications and word-of-mouth are likely to decrease this risk substantially.

3 The Demand Side

EPC trace data will prove interesting to many parties. A collection of possible uses is offered as an IBM patent application from a year as early as 2001 [14] that was made more visible later by [2].[7] The following is directly cited from [14]:

> *In another embodiment, instead of determining the exact identity of the person, some characteristics such as demographic (e.g., age, race, sex, etc.) may be determined based on certain predetermined statistical information. For example, if items that are carried on the person are highly expensive name brands, e.g., Rolex watch, then the person may be classified in the upper-middle class income bracket. In another example, if the items that are carried on the person are "female" items typically associated with women, e.g., a purse, scarf, pantyhose, then the gender can be determined as female. [...]* (p. 2) *When a person enters a retail store, a shopping mall, an airport, a train station, a train, or any location where a person can roam, a RFID-Tag scanner located therein scans all identifiable RFID-Tags carried on the person [...]* (p. 3)

This patent application is at least a clear hint that the authors did foresee a ubiquity of tags and readers and their implications. In the following we try to give a more detailed view on the motivations to collect and buy traces.

[7] Note the subtle irony of the fact that today IBM is also in the market for RFID privacy solutions, see [17].

3.1 Governmental Demand

For government agencies it will often be more convenient to buy or access personalized or raw traces from private companies, than to invest into additional reader infrastructures that cover enough areas for permanent surveillance.

Today, in the US, government agencies often buy personal data from profile brokers like ChoicePoint.[8]

As it saves expensive investments by the public hand, this trend will extend to traces and will at least continue until enough public readers (e.g., for ticketing or traffic monitoring and billing) are ubiquitously in place, which can then be used for secondary harvesting purposes. We give some examples for potential interest of governments in traces.

Customs and Tax Collection

Ownership of goods, their transfer and movement patterns will arouse strong interest by customs that can now track imported and exported goods. Likewise, tax collection for luxury items will be made easier by tracking items and their owners. Simply the anticipation of this possibility might reduce delicts and misdeameanour.

Law Enforcement

The police will have a high interest in traces. Areas where they will be extraordinary useful are crime investigations or forensics, e.g., to answer questions like where have these particular boots been sighted within the last month. Monitoring and remote surveillance of criminals or suspects (whether guilty or innocent) will be made much easier. Further, new ideas like supporting civil desaster recovery plans such as in case of epidemics may become thinkable once experience with the new technology has grown, and trace databases have become sufficiently large and accurate.

Intelligence Agencies

Even if they probably might already have access to equivalent information that traces could deliver, such traces could be used as confirming evidence to reduce uncertainties. Live traces could support other forms of surveillance, social (e.g., terrorist) networks could be analysed more easily.

3.2 Company Demand

The private sector will develop a huge demand for traces. Even companies that gather their own supply of traces will have reason to buy more (or trade) for data completion, integration and refreshment.

Personalization

There are empirically verified economic benefits for companies to personalize their goods or services [23][21]. Especially personalization and recommendation systems will be highly applicable to in-shop RFID applications and will also increase demand for trace data in E-Commerce. Some benefits of personalization are (cf. [25]) the ability to turn casual browsers into buyers, the

[8] `http://www.epic.org/privacy/choicepoint/`

potential of cross-sells by recommending matching items to something already owned or selected by the customer, and increased customer loyalty.

Personalized insurances[9] will stimulate a huge demand for traces by insurance companies to study a person's whereabouts, movements and consumer habits.

Traces will also enhance the effectiveness of credit scoring enormously by providing detailed insights into the subjects' possessions and purchases.

Price Discrimination and Direct Marketing

Price discrimination has been identified as an important driver for privacy erosion on the Internet [20]. To maximize profit a customer should ideally pay the maximum amount that is acceptable to her. To be able to charge different customers different prices for the same services or goods, data is needed to estimate their willingness to pay. Personalization of shopping sites, click tracing and other measures have been applied for that reason on the Internet. Traces will be a marvellous new source of relevant information even in the real world, answering a long list of interesting questions:

Product-related: Which products are likely to be bought together? What EPCs do people carry who (do not) enter my store or buy my products?

Customer-related: How do customers move in my store / all day? What are the characteristics of my customers? What do they already own? What are their budgets? How do customers and non-customers differ? Which products should I offer additionally to a customer?

If each customer in a shopping mall is a walking list of the items she has bought, like a personalized customer she is a much more convenient target for product placement and aggressive direct marketing of all sorts.

Industrial Espionage

Players in many industries will be interested in e.g. the transparency of a competitor's supply chain or simply lists of items or persons that enter their buildings. Also, less aggressive business intelligence can make use of EPC traces as well: "Which competitors have made which deals with which of my prospects?" Traces could offer all this information much easier than possible today.

3.3 Individual Demand

Individuals will also be interested in buying personalized traces. This might be to quench the natural human curiosity, or for more sinister activities like blackmailing or spying on neighbours, relatives or co-workers.

3.4 Research Demand

Finally we foresee a huge demand in raw data to support scientific research. Examples include research in epidemics, which could benefit from individual movement data, migration and mobility research and social sciences in general.

[9] An example today is "Pay as you drive": `http://www-1.ibm.com/services/ondemand/norwichunion.html`

4 Privacy Laws and EPC Traces

In some countries there are laws that restrict the ways in which governments, corporations, and citizens are allowed to collect, store, distribute and make use of personal information. However, the EPC traces under scrutiny here are not personalized *per se*. They can only provide privacy-relevant information by use of data mining techniques and the consultation of further data sources (such as a data warehouse that may contain a mapping from EPCs to customer IDs). Every single act of trace harvesting that is noticed and brought before a court will likely not have caused any damage, as the data is only valuable in large amounts. Therefore, it is unclear whether trace collection is covered by privacy laws at all. If it is not, trace harvesting and trading are perfectly legal.

4.1 The Question of Enforcability

Suppose the problem becomes apparent to the public and to policy makers at some point in the future, despite all marketing and lobbying efforts, and there will be a robust law prohibiting trace harvesting and trading. If the incentive structure is in favour of trade, an illegal market will emerge in replacement of the legal markets of traces, just like the markets for illegal drugs, botnets, or credit card numbers are perfectly healthy and operational despite (and sometimes due to) the laws against them. The only parameters affected by laws are the supply cost (i.e., the price of the trace harvesting and trading infrastructure), and the product quality (i.e., the trace density), but we do not expect the effects of any laws on these parameters to be very impressive.

There have been laws against spam and telemarketing in several countries for years now, but the industry is still prospering, because it is impossible to prove which of a number of legal bodies made a certain phone call or sent a certain e-mail. Networks of globally operating companies have emerged that are exceedingly cumbersome to hunt down and blame for any particular action or event. In contrast to spam activity, the effects of EPC trace collectors are both harder to notice for the individual and larger in number, both of which make the situation worse for the individual.

4.2 De-Anonymization

Laws usually introduce some notion of *how personalized* information is. For example, customer profiles that are aggregated to one tuple for every five households may be considered anonymized, or a certain volume of EPC trace information may be considered harmless. However, in practice even a single aggregated data source often contains enough information for a good data analyst to de-anonymize it to a large extent,[10] and a data pool of EPC traces and other corpora that is too enriched to be legally owned by one company can simply be split up between networks of companies, and the knowledge can be combined in a way that is certain to go unnoticed until aggregate data emerges in a completely different context.

[10] A much quoted figure is that 87% of all Americans are uniquely identifiable from ZIP code, birth date, and sex, so aggregation needs to destroy this information .

4.3 From Personal Data to Intellectual Property

There may even be laws such as intellectual property rights antagonizing privacy regulations. Once a market research agency has aggregated the collected trace data and thereby turned it into intellectual property, it is not only accessible to the wrong parties, but also harder to obtain by the affected individuals.

4.4 Explicit Consent where there is no Contract?

Explicit consent to collection of personal data can of course not be given if there is not even a contractual relation between the spied-upon end user and an independent trace collector operating in public spaces or in private spaces not owned by him. On the other hand, end users have to accept the terms imposed by the owner of a space they enter, for example video surveillance in shopping malls. By analogy, it seems likely that the owner of a shopping mall may – legally – decree that all goods sold on its premises are equipped with RFID tags, and that trace tracking is performed there.

5 Economic and Social Implications

Several authors have attended to the question of economic incentives for dealing with privacy. In understanding the reckless behaviour of individuals with their personal data, it is often claimed that people are irrational about the way they value privacy and are often willing to consent to the release of personal data even if privacy concerns exist [3]. Or, it is argued that privacy invading technologies succeed because "people are willing to give up their privacy for the benefits gained". This in turn is interpreted as a matter of "free-choice" [26]. Another approach is to interpret privacy as a trade-off [1].

We believe that these definitions are all short of covering the privacy concerns and economic implications that appear as a result of the markets of EPC traces. From an economic perspective an increase in information should be accompanied by a greater benefit resulting from it. This applies to competitors and consumers alike however, too much consumer information without any proper control mechanism could lead to negative externalities for the consumer despite the newly created positive externalities.

We have identified three main players in the market for EPC traces: the consumer, the collector of information and the end user. We assume governmental agencies to be part of the latter category. The interactions between the various groups are described below:

- Individual and Collector: Information in this relationship can be collected legally or illegally. Economic models refer to this arrangement as a trade-off whereby the individual provides personal information (disutility) in exchange for customized goods and services (utility).
- Collectors and End Users: In this stage, information that is collected in one market is utilized in a different market. The aim of the collector is to profit maximize as she knows the value of the traces being collected otherwise she would not participate. Further end users may be unwilling to give details

about the information they require. There are possible contract problems that can occur due to price and information asymmetry.

- End Users and Individuals: The information utilized by the end user is to generate better products and services for the individual based on the learning that was facilitated through better modes of generating data.

In principle, such technologies are geared at improving welfare by reducing transaction costs. They shift the production possibility curve and generate value. A better understanding of demand results in higher efficiency gains. It allows price discrimination which covers a larger spectrum of the market, i.e., more people have access to certain products and services, e.g., healthcare for the poor, differing prices for transport, movie tickets, etc. There is but more to what the relations between the three players listed above can imply. In the following we take a closer look at some of the issues that may arise in the EPC markets between these players.

5.1 Relationship between Collectors and End Users: Quality and Reliability of Aggregated Data

During the initial phases of the market it may not be clear how suppliers of EPC traces can best select the locations of their readers and find the appriopriate customers for their trace collections or aggregated data. Hence, the market may first have to be initialized by contracted suppliers with clear target EPCs and locations. Such contracts will be based on trust and accountability since customers of data will not only be collecting data, but revealing information on their data of interest to suppliers of data. The prerequisites to the establishment of such trusted contracts is beyond the scope of this paper. In another scenario, early locations may be limited to expected loci of interest like main squares, shopping centers, airports etc. In both cases, a wider market may first emerge once the demand for EPC traces and personalized data is better defined and a need for location-diversity is articulated by the customers.

The market of EPC traces will be vulnerable to malicious actors who sell fake traces. Further, competitive partners may be interested in injecting false data in order to effect the analyses of their competitors. Fake data do not "trace" volitional events, meaningful "market-baskets", or clusters, hence divergences between "trace based models of reality" and "reality" itself my occur. This may lead actors utilizing the EPC traces to false analyses and decisions, and may have economic repercussions. In addition, it will be difficult to determine where the fake data originates from.

When fake data is available and detected, quality and reliability of data may become an important aspect of the market. Buyers of trace collections or aggregated data may demand some sort of data authentication. This may establish a hierarchy of trusted and untrusted data providers. Security mechanisms and reputation models may become an integral part of guaranteeing quality and reliability in the market. Mechanisms to establish quality guarantees will effect the prices of trace collections and resulting data.

5.2 Relationship between Individuals and Collectors: The absence of the Opt-out Option

With RFID tags moving into indispensible daily products, it may become impossible to make a conscious and "free" choice on the use of tags and the resulting release

of EPC traces. Since unprotected tags may be very small in size, it may be difficult for individuals to recognize their presence. Some may see this as an opportunity to offer "untagged objects" to those individuals with privacy concerns. In such a case, it may be difficult or inconvenient for individuals to verify the absence of tags. Further, depending on the development in the privacy market, "untagged-objects" may be accessible to an informed, privacy-aware elite, just as PETs are in the current state of the internet. Nevertheless, it will be difficult for individuals to opt-out of the tags, if at all possible.

Hence, many of the conveniences through the use of EPC traces come at the price of personal anonymity/privacy. The market for EPC traces is a complicated one in that we see the occurrence of both positive and negative externalities. The latter gives rise to market failure as the collector and end user of the information do not internalize the misuse of the individual's personal information. When people responsible for privacy are not those who will suffer from indiscriminate use of personal information, there will always be a problem for the individual. This could indicate that the trade-off mentioned above in the defintion of relationship between individuals and collectors does not really exist.

Further, those individuals who value their privacy may prefer to opt out of transactions in the market completely. On the Internet, where opting out is an option, the authors of [1] mention that individuals concerned in privacy may "chose not to purchase online, or to purchase less, meaning that a latent, potentially large market demand remains unsatisfied". In a world where opting out of tags is not possible, it is difficult to say how the individuals may choose to behvave.

5.3 Relationship between End-Users and Individuals: Profiling individuals and the masses

The collection of considerable amounts of individualized information, regardless of whether they are personalized, makes possible what in Foucaultian terms is called a "panoptic surveillance".[9] Observation of individuals across a population allows the modeling of norms in that population. As a result, those who do not fit the the norms may be acted upon with gratification or punishment. In economics, this can result in the use of demographic traces and other individual information to structure the world by creating categories of ideal consumers, markets, products and places.[4]

Through price discrimination and re-distribution results many individuals lose out as they are no longer attractive or must pay higher prices. In the case of social goods, this form of re-distribution is not efficient and leads to dead weight loss in terms of non consumers. The lines of such dead weight loss may run along the lines of social discrimination which is inherent in the categorization and ranking of particular consumers by their potential value [4]. For example, the authors of [22] show that the development of geodemographic systems based on detailed information on individuals and their movements have effected the development of regions. They warn against the use of such data to discriminate against areas in which individuals with profiles that don't match the interests of well-capitalized, private corporate and/or state interests reside. The behaviour of individuals in the presence of complex forms of redlining as a result of trace analysis may become a problem in future economies.

Last, a study by Katz and Hermalin (2004) indicates that perhaps a ban on certain information rather than control rights of information may be better for certain

markets which exist based on imperfect information, i.e., insurance markets. A distinguishing feature of the EPC traces market, however, seems to be that information is collected unintentionally and covertly for an intentional purpose. Therefore, if we employ the Katz-Hermalin argument, the only option is to ban the technology.

Having said that, there seems to be both a demand and supply for EPC traces albeit compromising the interests of individuals. The solution seems to go back to Coase who suggested that institutional measures must be ensured in order compensate the negative externalities involved in such technologies. Perhaps technical solutions should be further explored as prevention is a lot easier than finding a cure.

6 Conclusions

Under the very reasonable assumptions that (1) readers will become cheap and small and that (2) reading RFID tags cannot be noticed or stopped in general, we have argued that there is a strong incentive to aggregate and subsequently trade traces of EPCs. We have assessed the nature of the business models involved and listed a number of potential buyers for possible trace-based products. Further, we have scratched a few legal issues and implications.

During the writing of this paper, we have started to develop cost functions and expected benefits for both the supply and the demand side, using existing and projected reader technology, well-established data mining techniques, and realistic application scenarios as a foundation to data which is at least as hard as the assumptions we make, and that can be easily adjusted each time the assumptions are revised. Designing such a model has turned out to be an ambitious task, but it would allow us to project the size of the market and coverage and availability of data and give a more accurate understanding of the state of privacy we should expect for the future.

6.1 Acknowledgements

We would like to thank Michael Klafft for discussions of earlier versions of this paper.

References

[1] Alessandro Acquisti. Privacy and security of personal information: Economic incentives and technological solutions. In J. Camp and S. Lewis, editors, *Economics of Information Security*. Springer Verlag, 2004.

[2] Katherine Albrecht and Liz McIntyre. *Spychips*. Nelson Current, 2005.

[3] Jens Grossklags Bettina Berendt and Sarah Spiekermann. Eprivacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *In 3rd ACM Conference on Electronic Commerce - EC '01*, pages 38–47, 2001.

[4] Phillips David J. Privacy policy and PETs: The influence of policy regimes on the development and social implications of privacy enhancing technologies. *New Media Society*, 6(6):691–706, 2004.

[5] Saso Dzeroski. Multi-Relational Data Mining: An Introduction. *KDD Explorations*, July 2003.

[6] EPCglobal. EPC Tag Data Standards Version 1.1, 2004.

[7] Benjamin Fabian, Oliver Günther, and Sarah Spiekermann. Security analysis of the Object Name Service. In *Proceedings of the 1st International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing (SecPerU 2005), in conjunction with IEEE ICPS 2005, Santorini*, pages 71–76, 2005.

[8] Christian Floerkemeier, Roland Schneider, and Marc Langheinrich. Scanning with a Purpose – Supporting the Fair Information Principles in RFID protocols. In Hitomi Murakami, Hideyuki Nakashima, Hideyuki Tokuda, and Michiaki Yasumura, editors, *Ubiquitious Computing Systems. Revised Selected Papers from the 2nd International Symposium on Ubiquitous Computing Systems (UCS 2004), November 8-9, 2004, Tokyo, Japan*, volume 3598 of *Lecture Notes in Computer Science*, Berlin, Germany, June 2005. Springer-Verlag.

[9] Michel Foucault. *Discipline and Punish*. New York: Vintage Books, 1979.

[10] Simson Garfinkel, Ari Juels, and Ravi Pappu. RFID privacy: An overview of problems and proposed solutions. *IEEE Security and Privacy*, 3(3):34–43, May-June 2005.

[11] Simson Garfinkel and Beth Rosenberg, editors. *RFID Applications, Security, and Privacy*. Addison-Wesley, 2005.

[12] Simson L. Garfinkel. Adopting Fair Information Practices to Low Cost RFID Systems. 2002.

[13] Sumi Helal, William Mann, Hicham El-Zabadani, Jeffrey King, Youssef Kaddoura, and Erwin Jansen. The Gator Tech Smart House: a programmable pervasive space. *IEEE Computer Magazine*, pages 50–60, March 2005.

[14] John R. Hind, James M. Mathewson, and Marcia L. Peters. Identification and tracking of persons using RFID-tagged items. *US Patent Application*, (20020165758), 2001.

[15] Sozo Inoue, Shin'ichi Konomi, and Hiroto Yasuura. Privacy in the Digitally Named World with RFID Tags. In Workshop on Socially-informed Design of Privacy-enhancing Solutions in Ubiquitous Computing *UBICOMP'02*.

[16] Ari Juels. RFID security and privacy: A research survey. `http://www.rsasecurity.com/rsalabs/staff/bios/ajuels/publications/pdfs/rfid_survey_28_09_05.pdf`, September 2005.

[17] Günter Karjoth and Paul Moskowitz. Disabling RFID tags with visible confirmation: Clipped tags are silenced. In *Workshop on Privacy in the Electronic Society – WPES*, Alexandria, Virginia, USA, November 2005. ACM, ACM Press.

[18] Lin Liao, Dieter Fox, and Henry A. Kautz. Learning and Inferring Transportation Routines. In *Proc. AAAI*, pages 348–353, 2004.

[19] David Molnar, Andrea Soppera, and David Wagner. Privacy for RFID through trusted computing. In *Workshop on Privacy in the Electronic Society – WPES*, Alexandria, Virginia, USA, November 2005. ACM, ACM Press.

[20] Andrew Odlyzko. Privacy, economics, and price discrimination on the internet. In *Proceedings of the 5th international conference on Electronic commerce (ICEC '03)*, pages 355–366, New York, 2003. ACM Press.

[21] Don Peppers, Martha Rogers, and Bob Dorf. *The One to One Fieldbook*. Capstone Publishing Ltd, 1999.

[22] D. Phillips and M. Curry. *Privacy and the phenetic urge: Geodemographics and the changing spatiality of local practice.* London and New York: Routledge., 2002.

[23] B. Joseph Pine II, Bart Victor, and Andrew C. Boynton. Making mass customization work. *Harvard Business Review*, September-October 1993.

[24] Melanie Rieback, Bruno Crispo, and Andrew Tanenbaum. RFID guardian: A battery-powered mobile device for RFID privacy management. In Colin Boyd and Juan Manuel González Nieto, editors, *Australasian Conference on Information Security and Privacy – ACISP'05*, volume 3574 of *Lecture Notes in Computer Science*, pages 184–194, Brisbane, Australia, July 2005. Springer-Verlag.

[25] J. Ben Schafer, Joseph A. Konstan, and John Riedi. Recommender systems in e-commerce. In *ACM Conference on Electronic Commerce*, pages 158–166, 1999.

[26] A. Shostack. 'people won't pay for privacy,' reconsidered. 2003.

[27] Frank Stajano. RFID is X-ray vision. *Communications of the ACM*, 48(9):31–33, 2005.

 At last week's RFID Track and Trace Health Care Summit, there was a great deal of focus on what state governments and the FDA will require the pharmaceutical industry to do, and less emphasis on the business value of RFID. - Page 1.Â The role of electronic track and trace will be critical, he told attendees. The U.S. Congress requires the FDA to publish guidelines in March 2010, and according to Shuren, there is continued activity in Congress that could require the agency to take a more proactive role in establishing track-and-trace requirements. Paul Rudolf, a former FDA official and now a senior health-care advisor for law firm Arnold & Porter LLP, said he believes other states will follow California's lead and essentially enact pedigree laws with a similar timetable.