Emerging Markets for RFID Traces

Matthias Bauer (`matthiasb@acm.org`)
*

Benjamin Fabian (`bfabian@wiwi.hu-berlin.de`)
Matthias Fischmann (`fis@wiwi.hu-berlin.de`)
Seda Gürses (`seda@wiwi.hu-berlin.de`)
†

Abstract

RFID tags are held to become ubiquitous in logistics in the near future, and item-level tagging will pave the way for Ubiquitous Computing, for example in application fields like smart homes. Our paper addresses the value and the production cost of information that can be gathered by observing these tags over time and different locations. We argue that RFID technology will induce a thriving market for such information, resulting in easy data access for analysts to infer business intelligence and individual profiles of unusually high detail.

Understanding these information markets is important for many reasons: They represent new business opportunities, and market players need to be aware of their roles in these markets. Policy makers need to confirm that the market structure will not negatively affect overall welfare. Finally, though we are not addressing the complex issue of privacy, we are convinced that market forces will have a significant impact on the effectiveness of deployed security enhancements to RFID technology.

In this paper we take a few first steps into a relatively new field of economic research and conclude with a list of research problems that promise deeper insights into the matter.

1 Introduction

RFID tags are widely held to become ubiquitous in our future. They are already essential in logistics [11], and item-level tagging will pave the way for Ubiquitous Computing applications such as so-called smart homes [12]. Our paper is based

---

*Shoestring Foundation, Erlangen
†Institute of Information Systems, Humboldt-University, Berlin

on a still hypothetical (but widely accepted) Ubiquitous Computing scenario where goods are tagged on a common basis, and aims to demonstrate that the incentives in this scenario will induce thriving markets for information of the form "serial number i was at location l at time t". These markets will not only mediate between the tag wearing individuals and tag reader operators, but also, and more importantly, among reader operators, data mining service providers, profile dealers and data customers after the information has been harvested.

The main and currently most influential industry standards for tags, readers and number schemes are set by the industry consortium EPCglobal. Conforming EPC tags transmit an *Electronic Product Code (EPC)*, a data structure that has a globally unique value on each tag (see Appendix B). Where applicable, we use the terminology established by EPCglobal, our arguments, however, do not depend in general on any details in the standards. The central assumptions we use are made explicit in the following.

EPC tags for the mass market are small, short-ranged, and extremely cheap RFID devices.[1] In particular, a tag cannot keep track of who is reading it and when it is read, or selectively deny reading. Standards for more capable tags have been proposed that can selectively block read-outs or even communicate with trusted readers only over encrypted radio links.

Though the EPC Class 1 chip specification includes a kill password to deactivate a tag permanently, this would prevent any after sale application, and the usability of these or comparable features is still unclear (see Section 4.3.1). We claim that collection and pervasive distribution of tag data is essential to the usefulness of RFID technology. Thus our arguments about the emergence of trace markets would still remain basically valid even in the presence of powerful and usable security enhancements.

The envisioned ubiquity of tags, together with corresponding new (as well as existing) databases and sophisticated data mining techniques, promises an enormous amount of information on the movements of goods and individuals, and on correlations between entities and objects.

However, very little work has been done on getting accurate or even plausible estimates on the value of this information, the cost of harvesting it, or on finding a model that describes what will happen if this scenario becomes reality. To confuse the situation further, there are statements from industry leaders to the effect that interest in this information is negligible,[2] despite efforts to create test beds that facilitate access to it.[3]

On the other hand, understanding the economics of EPC traces is crucial for

---

[1] Reading ranges depend on frequency, antenna, surroundings and energy. Vendors are aiming at costs of less than 0.05$ per tag for the mass market.

[2] See the statement of the METRO Group during the consultation process by the EU Art. 29 Data Protection Working Party: http://ec.europa.eu/justice_home/fsj/privacy/workinggroup/consultations/rfid_en.htm

[3] METRO Future Store, http://www.future-store.org/

Figure 1: Players in EPC Trace Markets.

many reasons. Policy-makers need to address the question whether there is need for regulation. Market players want to know the ground on which they compete. Finally, in order to estimate the impact of Ubiquitous Computing technology on the privacy of individuals, it is necessary to understand the motivations and incentives of the parties that operate it.

The goal of our paper is to give a first description of EPC trace markets, identify some of their most obvious characteristics, and lay the ground for future research.

The paper is structured as follows: In Section 2 we give a brief overview of our terminology, the market players, and the structure of goods. In Sections 3 and 4 we informally assess the demand and the supply side, respectively. In Section 5 we compare the emerging EPC trace markets to the markets for online traces that exist today, drawing some conclusions on differences in interpreting information and on market risk. In Section 6, we present further economic research topics that need to be addressed in order to develop a deeper understanding of EPC trace markets.

2 Trace Markets

The raw product of EPC trace markets is a database of entries of the form (i, l, t): i is an identifier (in our case: an EPC); t is a time value (relative or absolute); and l is a location identifier (a GPS coordinate tuple, site of the supermarket, a reader number, IP address, or any useful and consistent description of the reader's context). 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 or PDA.

We distinguish three main categories of market players (see Figure 1):

- The *Individuals (Data Subjects)*. Entities that possess tags and generate traces. This category includes private citizens and employees of companies.

- The *Data Consumers*. The end users of trace data who require market profiles or information on individuals for various reasons. (See Section 3.)

- The *Data Suppliers*. Suppliers of trace data may be collectors who operate tag readers and feed databases, or aggregators who provide services by

trading, pooling, and data mining such databases. (See Section 4.)

The relationships and interactions between these categories are complex for various reasons.

Individuals and Data Suppliers. Information in this relationship can be collected legally or illegally. There are economic models in which individuals face a trade-off, providing personal information (disutility) in exchange for customized goods and services (utility). Calibrating the involved utility functions is an ambitious goal.

Data Consumers and Individuals. Being better informed about the market, the data consumers are able to offer better products and services for the individuals. The information available from trace databases can greatly facilitate price discrimination. The reliability of data and the appropriate degree of price discrimination based on this is to be studied.

Data Suppliers and Data Consumers. The players in this relationship no longer consider the EPC traces to be personal information. They rather regard them as tradeable information goods that can be sold to data consumers, who may then use the data to improve their own competitiveness in another, independent market.

The authors of [3][22][1] address the first two relationships. In this paper we focus on the relationship between data suppliers and data consumers, whose interests disregard those of individuals.

3   The Demand Side

There are indications that RFID traces in general and EPC traces in particular will prove to be instrumental to many parties. For example, a collection of possible uses is offered in an IBM patent application from as early as 2001 [13] (pointed out later by [2]):

> *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, panty hose, 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 gives anecdotal evidence that some experts did foresee the potential usefulness of gathering quality information through RFID traces.[4] In this section we try to give a more detailed view on the motivations of potential trace consumers. We consider companies, governments, individuals, and research organizations.

3.1 Companies

There are many reasons for the private sector to develop a substantial demand for traces. Even companies that gather their own supply of traces will have reason to buy from or pool with other companies for data completion, integration and refreshment.

Personalization. There are empirically verified economic benefits for companies to personalize their goods or services [20]. This has influenced industry best-practices [19]. Especially personalization and recommendation systems will be highly pertinent to in-shop or home RFID applications (B2C) and will also increase demand for trace data in E-Commerce. Some benefits of personalization are (cf. [21]) the ability to turn casual browsers into buyers, the potential of cross-sells by recommending matching items to something already owned or selected by the customer, increased customer loyalty and better customer relationship management.

Personalized insurances[5] will stimulate a huge demand for traces by insurance companies to study a person's whereabouts, movements and consumption habits [4]. Traces will also enhance the effectiveness of credit scoring enormously by providing detailed insights into the subjects' possessions and income. An indicator is today's market for corresponding information (see Appendix A).

Price Discrimination and Direct Marketing. Price discrimination has been identified as an important driver for privacy erosion on the Internet [18]. To maximize profit a customer should ideally pay the maximum amount that is acceptable to her. In order to charge different customers different prices for the same services or goods, data is needed to estimate their willingness to pay. Data generated through personalization of shopping sites, click tracing and other measures used on the Internet are conducive to such analysis [5].

Traces will be a new source of relevant information that also pertain to the physical world, answering interesting product-related and customer-related questions: Which products are likely to be bought together? What EPCs do people carry who do (do not) enter my store or buy my products? How do customers move in

---

[4]Interestingly, IBM has also entered the market for RFID privacy solutions later [16].

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

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 additional products should I offer to a customer?

If each customer in a shopping mall is recognized by the IT infrastructure as a walking list of the items she bought (nearly equivalent to a fully personalized customer), she is a much more convenient target for product placement and direct marketing strategies.

Industrial Espionage. Players in many industries will be tempted, if not actively interested, in the possibility of inspecting 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 deals with which of my prospects? What kind of deals are these?" Traces can make this information more accessible than it is today.

Intensified Competition for Information. In a competitive environment where each individual player can decide to increase her data pool in order to increase her advantage over other players, an arms race in business intelligence can emerge. The following game can be used to model this effect. Let the information held by company i in the current round be I_i, let c, v be monotonously increasing functions describing the cost and value of that information, respectively, and let the utility U_i of company i be[6]

$$U_i = \frac{v(I_i)}{\mathbf{avg}_{j \neq i} v(I_j)} - c(I_i).$$

Then, if company i increases its information in the next round to $I_i^* = I_i + \delta_i$, company j is worse off unless it increases $I_j^* = I_j + \delta_j$ such that

$$U_j^* = \frac{v(I_j^*)}{\mathbf{avg}_{k \neq j} v(I_k^*)} - c(I_j^*) \geq U_j.$$

This holds assuming that all other players do not change their information levels. However, as now both companies i and j have increased theirs, all other players have even greater pressure to do so as well in the next round.

3.2   Governments

For government agencies it will often be more convenient to accumulate raw or personalized traces from private companies, rather than to invest into additional reader infrastructures that cover sufficient area for permanent surveillance.

Today, in the US, government agencies often buy personal data from profile brokers like ChoicePoint.[7] We expect this trend to extend to the EPC traces at

---

[6]A more accurate utility function can no doubt be developed. But as long as U_i increases with I_i and decreases with $I_j, j \neq i$, our argument holds.

[7]References gathered by EPIC at http://www.epic.org/privacy/choicepoint/

least until enough public readers are installed (e.g. for ticketing, traffic monitoring, billing and building security). We list some of the potential utilities of EPC traces for governments below.

Customs and Tax Collection. Ownership of goods, their transfer and movement patterns are very interesting to custom services that could now track imported and exported goods. Likewise, tax collection for luxury items will be made easier by tracking items and their owners. Simply the threat of this possibility might reduce delicts and misdemeanor.

Law Enforcement. The police will have a high interest in traces, as they will prove extraordinarily useful in forensics. They could be used to address questions like "Where have these particular boots been sighted within the last month?". Monitoring and remote surveillance of criminals or suspects will be made much easier. Further, ideas like supporting civil disaster recovery plans through new technologies, e.g. in case of epidemics, may become possible once trace databases have become sufficiently large and accurate.

Intelligence Agencies. Even if they already have access to equivalent information that traces could deliver, 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 analyzed more easily. Nevertheless, the challenge of false positives and false negatives will have to be tackled even more seriously with the increase of EPC traces.

3.3 Individuals

Individuals will also be interested in using trace databases. This might be to quench natural human curiosity, or for more sinister activities like blackmailing or spying on neighbors, relatives or co-workers. On the other hand, applications for child care as well as care for elderly persons make use of Ubiquitous Computing data, and are appreciated by their users for improving their quality of life considerably.

3.4 Researchers

Finally we foresee a substantial demand in raw data to support scientific research. Examples include economics (e.g. improving research on trade), research in epidemics, migration and mobility research, and social sciences in general.

4 The Supply Side

The structure of the supply side of a market decides to what extent anybody will engage in production. In order to understand this, we look at the reader locations where data is physically harvested, and the globally connected IT infrastructure where it is exchanged and processed. Then, we assess potential obstacles that may perturb the trace supplier's job, namely technical and legal countermeasures.

4.1 Local EPC Traces

In Ubiquitous Computing environments like smart buildings RFID will be a fundamental technology to discriminate contexts. Applications like smart fridges, laundry machines and home medical advisors could use RFID and the EPC Network to provide inhabitants with new services. Extensive reader infrastructures will be installed for these or other primary purposes and may be used officially and with implicit consent for trace harvesting. For example, these technologies can be employed inside a shop at the cashiers, scanning and tracking the shop's EPCs (and those that the customers brought with them).

For service providers there lies plenty of opportunity for the secondary use of these collected traces, which will often be already associated with an individual (like in the case of smart homes) or easily linkable (in the case of shop customers). As the market for traces develops, the reader operation is likely to be outsourced to specialized service providers. These will serve a greater set of customers in multiple settings and have access to larger volumes of data. If their contracts permit them to do so, these specialized service providers can also sell this data to formerly uninvolved third parties.

There also could be rogue readers operating despite the dissent of the inhabitants of their environments. Mobile or fixed installations can be used to scan strategic areas within cities or buildings to gather traces without consent or notice. More secretively, there is the possibility to place passive sniffers[8] near official reading infrastructures to capture RFID communications and collect transmitted EPCs.

Finally, creative marketing projects might come up with gadgets like a portable music player that traces tags around its owner, producing more or less a complete profile of her and even 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. Such business models could be designed in the spirit of customer loyalty programs.

---

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

4.2   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 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 as a backbone to an *Internet of Things* and to Ubiquitous Computing environments.

The EPC Network (like every solution for EPC Information retrieval via the Internet) will constitute an excellent trace trading system and, possibly even to a larger extent than local RFID interaction, a global trace-collection mechanism. Every EPCIS can itself be a trace collection point, simply by log file analysis. Instead of the physical location an EPCIS collects the source IP of the query, and probably further identifiers and authentication tokens from the information protocol layer. Third parties cannot sniff this trace-containing traffic if it is encrypted.

But the look-up service *Object Naming Service (ONS)* is a weak spot in the proposed design [8]. ONS uses the structure of an EPC, which is transformed into a domain name, for query delegation and in functionality is akin to the Internet's Domain Name System (DNS). Every ONS server in the resolution chain is a potential trace collection point.[9] ONS is a clear text protocol and is about to inherit all the well-known security weaknesses of DNS. In the worst (or best, respectively) case, every IP node in the ONS resolution chain, every router or network analysis device in the path will be able to collect partial traces.

We anticipate that EPC information exchange over the Internet, and especially the EPC Network, will make traces even more widely available, in addition to those generated by local and direct skimming via RFID readers. Collection and data aggregation systems like intrusion detection systems that are in place today could be modified to harvest EPC trace data from the Internet. Combining these into large, global aggregation services that exist for, for example, distributed intrusion detection[10] today, or including them in search engines like *Google* or RFID-pendants of services like *Where's George*[11] might let lookups for specific EPC clusters succeed with high probability.

---

[9]The ONS specification as of today does not use the serial number for delegation, but leaves it as option for the future. In its current form, identifying and tracing rare item classes or clusters of items will be possible.

[10]DShield: `http://www.dshield.org`

[11]`http://www.wheresgeorge.com/`

4.3 Restricting Supply

In the following sections, we look at the strategies for restricting the unwanted use of EPC traces and study their possible repercussions for the trace trade markets.

4.3.1 RFID Protection: Restricting the Information Flow

If tags are ubiquitous, a lot of privacy issues arise [2][10][23]. Several security enhancements to control the flow of information have been proposed to maintain a certain level of privacy for individuals [10][15].[12]

In order to estimate the density of traces available for harvesting and trading, one needs to understand these security enhancements. In this section, we give a very brief overview of some existing data protection methods and identify problems that may negatively affect their efficiency in practice.

Tag readers are publicly available and will become cheap. Hence, anyone will be able to read all the tags located within a few meters distance and with minimal effort maintain traces of the area around her readers. Consequently, not only could operators of legitimate readers re-use the data for trading, but there could be reader operators that collect traces solely for the trace markets. Various enhancements to block or control read-outs have been proposed in order to confine the set of readers to those deemed legitimate by the individual wearing the tags. Although promising, these solutions face a number of obstacles.

Selective Deactivation or Blocking. This approach is inconvenient and (without strong authentication) highly problematic. If a chip needs to be disabled every time a hostile reader might be nearby, and re-activated every time a friendly reader is expected to read it, things will frequently go wrong. The user will forget to disable tags, or hostile readers will be installed close to friendly readers, and the user will have to choose between not profiting from the tag infrastructure and being exposed to a hostile read-out.

Symmetric Cryptography. Authenticating readers (possibly also the tags) and communicating over an encrypted channel can make blocking out untrusted readers both more convenient and more reliable. Symmetric cryptography can be used for both authenticity and encryption.

Since the storage capacities of a tag are extremely limited, only a small number of symmetric keys can be stored on a tag. All readers potentially communicating with a tag need to possess one of the keys stored on the tag. This establishes *trust clusters* among large numbers of readers and huge numbers of tags. Each message can only be shown to originate from a certain cluster, not from a certain

---

[12]A dedicated Web site for RFID security and privacy research is maintained by G. Avoine: `http://lasecwww.epfl.ch/~gavoine/rfid/`

device. Consequently, any adversarial trace collector with physical access to any tag or reader from a given trust cluster can retrieve the symmetric key for that cluster, pretend to be a trusted reader, and harvest the protected tags.

Asymmetric Cryptography. Whereas symmetric authentication and encryption in applications with many communication partners quickly exhaust the key management capabilities of the parties involved, asymmetric cryptography exceeds the computing resources of most RFID tags. Especially tags that have no batteries and need to be small and cheap enough to tag small and cheap goods are not suitable for the implementation of such mechanisms. Asymmetric cryptography is at best considered an option for tags with higher capacity used in production and logistics.

Trust. A more exotic option is to broadcast privacy policies from a PDA [9]. This PDA would ask all readers in scanning distance to ignore all sensitive EPCs, and law enforcement will provide the incentives for reader operators to obey the policies that are expressed. This approach is likely to fail in the trace application scenario for at least two reasons: First, economic incentives to get to the data are strong, as this paper demonstrates; second, with P3P, an existing system based on this idea has proven too inconvenient for the user to be widely deployed [14]. So in practice, neither opt-in nor opt-out are easy options.

Writing privacy policies is complex and cumbersome. More importantly though, content providers on the Web have the power to make it inconvenient for users to choose a restrictive privacy policy. Usually they can do this by configuring unnegotiable policies, leaving it to the user to either defer the use of a site, or to relax her own preferences. In the RFID future we envision, those players planning EPC-based services will be able to reiterate this practice, but then not only Web browsing behavior, but rather the activities of the user in the physical world will be affected.

Link between Company and EPC Network. The security of current EPC Network and ONS drafts is, to say the least, controversial (see Section 4.2 above). So even if the link between tag and reader would be perfectly secure, the information flowing through the global EPC infrastructure (of which readers are only the leaves) provides a potentially even richer field for trace collectors and brings new challenges to IT security.

4.3.2   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 raw 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).

Each single act of trace harvesting that is noticed and brought before a court will likely not have caused any damage on its own, as the data is only valuable in large amounts. Therefore, it is unclear whether trace collection is covered by privacy laws at all.

The Question of Enforceability. Suppose the problem becomes apparent to the public and to policy makers in the future and a law prohibiting trace harvesting and trading is passed (despite the economic advantages of collecting such information). If the incentive structure is in favor of trace trade, an illegal market will emerge in substitution of the legal markets of traces, just like the markets for illegal drugs, botnets, or credit card numbers are perfectly healthy and operational despite the laws against them. Laws can affect the cost of supply (i.e., the price of the trace harvesting and trading infrastructure) and the product quality (i.e., the trace density), but whether the impact on trace markets in particular will be high enough to stop trade is uncertain.

De-Anonymization. Privacy 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 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.[13] A data pool of 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.

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 paying third parties, but also harder to obtain by the affected individuals.

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 profiled individual and an independent trace collector operating in public spaces, or in private spaces not owned by him. On the other hand, 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 tracking is performed.

---

[13]A much quoted figure is that 87% of all Americans are uniquely identifiable from ZIP code, birth date, and sex, so aggregation needs to obfuscate this information [24].

5 EPC Traces vs. Web Traces

We will now take a closer look at an already existing information market that might prove insightful to EPC traces markets, although it seems to have been exposed to only sparse economic research: The market of Web Analytics.[14] Visitors of the World Wide Web produce traces in server logs that are often subject to complex business models. One special tracing method is the so called *Web Bugs*[15], nearly invisible linked images that allow tracking of users across different Web sites or via electronic mail. Among other things, they are used to verify how many users viewed an ad or read an e-mail. Consequently, despite privacy concerns, they constitute a popular basis for advertisement pricing and market research.

Web traces are similar to EPC traces, the EPC corresponds to a client IP (and often an additional session ID), the location information obtaining the form of a URL being accessed. We argue that EPC trace databases will be similarly successful, if not more so, as they mirror movements in the physical world. However, they differ in important details.

Data. EPC traces per se 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 on mornings, but never together in the afternoons).

Aggregation Technology. EPC traces bring about new algorithmic challenges: Needed are algorithms on spatio-temporal data of high resolution; data integration to obtain a semantic and thus actionable results; and algorithms that work on data with a rich relational structure.

Such algorithms have been developed over the past decades. 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 [17]. Advances in (multi-)relational data mining [6] are also substantial. Thus, if the supplier of EPC traces offers analysis options, the costs for additional intelligence are low.

The Passive Role of the User. EPC traces arise from ubiquitous contexts and not from limited interaction with software via an interface that can be configured by the user. The consequences are the occurrence of more traces, and in principle traces that have higher density than the corresponding traces collected on the Web. On the other hand, the interpretation of what an ID-time-location triple means may become even harder than with click traces. In the

---

[14]This market is changing at the moment, because a big player has entered the stage: Google now provides analysis of click traces for free (Google Analytics http://www.google.com/analytics). Such services are possible for EPC traces as well.

[15]EFF Definition: http://www.eff.org/Privacy/Marketing/web_bug.html

Web trace business, errors are sometimes intentionally fed into the databases by privacy-aware users. Another aspect of the passive role of the user in the collection of EPC traces is that she cannot effectively lie about the data that is generated.

The Problem of Coverage. The spatial entities that are being tracked are different, and so is the locus of control. EPC tracing also starts from locations, but these are usually public and exist independently of the objects that are being tracked. In the absence of universal tracking coverage, the collectors of traces cannot guarantee the collection of data that might be of interest to a given customer.

This will be mitigated by the following circumstances: First, the trace collector may be the owner of the space under surveillance. In this case, the situation is very similar to the Web, the market risk is low, and good coverage is affordable. Second, the trace collector may have entered into a contract to collect traces at specific locations. The risk that no interesting data are collected must be borne by either or both of the contract partners, and appropriate pricing models will need to be developed. Third, these bootstrapping problems do not exist if companies already gather data for purposes other than trading, yielding tradeable traces as a positive externality.

Summing up, the collection and analysis of EPC traces seem to imply only manageable additional business risks compared to Web tracing, but will encounter much higher demand, as there are more customers interested in tracing the physical than the virtual world.

6   Economic Research Problems

In the following we discuss some issues that may arise in the EPC markets between the data consumers and the data suppliers, and identify future research questions in modelling this relationship.

Web Trace Markets. What is the state of Web trace markets today, how much money is made? Similarities and differences to EPC trace markets need to be studied further.

Initialization of EPC Trace Markets. We introduced some initialization scenarios in Section 5. 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 appropriate customers for their trace collections or aggregated data.

Information Asymmetry. Contracts between trace collectors and trace consumers will include information on target EPCs and locations. Customers of data will reveal their specific interest in the process, while the supplier won't give data away indiscriminately. This could cause contracting problems and

may result in information asymmetries that may lead to market failures. The implementation of design mechanisms to avoid such a market failure is a topic of future research.

Quality Assurance. The market of EPC traces will be vulnerable to malicious actors who sell fake traces. Competitive partners may be interested in injecting false data in order to effect the analyses of their competitors. This may lead actors utilizing the traces to false analyses and decisions, and may have economic repercussions.

7 Conclusions

Under the assumptions that networked RFID technology and usage of the Electronic Product Code in general become ubiquitous, 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 investigated potential buyers of trace-based products. A closer look at the supply side brought us to the preliminary conclusion that despite potential technical and legal barriers trace markets will emerge.

Comparing the trace market to Web Bugs has given indicators that business risk for data providers and market initialization will be manageable. Further, we have sketched some topics for further economic research concerning the relation of data consumers and providers.

Evaluating our results in the broader context of society, we come to the conclusion that ultimately a balance between two extreme states needs to be found: One extreme would be a state in which only secure RFID technology is introduced that doesn't produce much information on individuals or goods. Given the current technological developments this state is impractical and economically ineffectual. In the other extreme state security is not an issue, information flows perfectly free (and hopefully for the benefit of both individuals and the corporate world), but privacy is drastically eroded by the emerging information markets.

Crucial questions are: Where do we, as individuals, want to be between these two possible states in our daily lives, potentially *despite* the market forces? And will scalable and usable security and privacy enhancements to RFID and EPC systems be implemented to get us there?

8   Acknowledgments

References

[1] Alessandro Acquisti and Hal R. Varian. Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381, 2005.

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

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

[4] Jürgen Bohn, Vlad Coroama, Marc Langheinrich, Friedemann Mattern, and Michael Rohs. Living in a world of Smart Everyday Objects – Social, economic, and ethical implications. *Journal of Human and Ecological Risk Assessment*, 10(5):763–786, Oktober 2004.

[5] Lorrie Faith Cranor. 'I didn't buy it for myself' – Privacy and E-Commerce personalization. In *Proceedings of the 2003 ACM workshop on Privacy in the electronic society (WPES03)*, pages 111–117, 2003.

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

[7] EPCglobal. EPC Tag Data Standards Version 1.3, 2006.

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

[9] Christian Floerkemeier, Roland Schneider, and Marc Langheinrich. Scanning with a Purpose – Supporting the Fair Information Principles in RFID protocols. In *Ubiquitious Computing Systems. Revised Selected Papers from the 2nd International Symposium on Ubiquitous Computing Systems (UCS 2004), November 8-9, 2004, Tokyo, Japan*. Springer, 2005.

[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] 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.

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

[14] Harry Hochheiser. The platform for privacy preference as a social protocol: An examination within the U.S. policy context. *ACM Transactions on Internet Technology*, 2(4):276–306, 2002.

[15] Ari Juels. RFID security and privacy: A research survey. *IEEE Journal on Selected Areas in Communications*, 24(2):381–394, February 2006.

[16] 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.

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

[18] 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.

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

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

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

[22] A. Shostack. Paying for Privacy: Consumers & Infrastructures. In *2nd Annual Workshop Economics and Information Security, University of Maryland*, 2003.

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

[24] Latanya Sweeney. On standards of privacy of individually identifiable health information. http://privacy.cs.cmu.edu/dataprivacy/HIPAA/HIPAAcomments.html, 2002.

A   Appendix: Profiles as Products, Today

RFID and EPC traces will increase the granularity of profiling individuals, creating new paths for data mining to extract high level information. For some data there already exists a market today.

The following is a partial price list (May, 2006) cited from the SWIPE project (`http://www.turbulence.org/Works/swipe/`). The prices are collected from commercial warehouses and profile brokers (Accurint, Aristotle, ChoicePoint, ChoiceTrust, DocuSearch, Experian, KnowX, Merlin Data, and Pallorium). All prices in $.

General	
Address	0.50
Zip Code	0.50
Past Addresses	9.95
Date of Birth	2.00
Marriage	7.95
Divorce	7.95
Education	12.00
Employment	13.00
Published Phone #	0.25
Unpub. Phone #	17.50
Cellular Phone #	10.00
Past Phone #s	0.50
Relatives	3.00
Neighbors	0.25
Registered URL/Domain Name	0.25
Soc. Security #	8.00

Financial	
Credit	9.00
Real Estate	1.50
Bankruptcy	26.50
Worker's Compensation	18.00
Assets	6.95
Assets Seized	2.95
Shareholder	1.50
Executive Affiliation	0.50
Own Aircraft	1.50
Own Boat	1.50
Own Vehicle	0.75
Own Business	9.95

License Info	
Driver Lic. Info	3.00
Motor Veh. Reg.	3.00
List of Vehicles	0.70
Accident Reg.	1.00
Aircraft Lic.	1.50
Drug Enforcement Admin. (DEA) Lic.	0.25
Hunt & Fish Lic.	0.25
Professional Lic.	0.75
Industry Accreditation	16.00
Merchant Vessel	0.25
Concealed Weapon	0.25
Firearms Lic.	0.25

Legal	
Lawsuits	2.95
Felony	16.00
Misdemeanor	9.00
Sex Offender	13.00

B Appendix: Electronic Product Code (EPC)

The most prevalent kind of EPC (SGTIN-96) has a length of 96 bit and corresponds to a SGTIN (Serialized Global Trade Identification Number). The main structure of this EPC is as follows [7] (see Figure 2):

Figure 2: Electronic Product Code (SGTIN-96 EPC)

The Header defines the kind of number that is encoded (here SGTIN-96). Filter Value is a rough object classification (e.g., retail item). Partition determines the exact boundary between the two following values. Company Prefix (former EPC Manager) determines which company issued this EPC (usually the manufacturer of the corresponding item). The Item Reference (Object Class) determines the exact category that the tagged object belongs to, and the Serial Number identifies a particular item within the same object class.

Note that this serial number enriches the information carried by an EPC tag significantly compared to a bar code, making distinction between individual items of the same kind possible.