Agricultural Supply Chains: Track and Trace for Improved Food Safety

E.W. Schuster
Field Intelligence Lab and Data Center Program
Laboratory for Manufacturing and Productivity
Massachusetts Institute of Technology
Cambridge, MA 02139

Keywords: food safety, supply chain, RFID, M Language, MIT

Abstract

Safety has emerged as an issue for the fresh fruit and vegetable industry with reports of several food poisoning outbreaks traced to E. Coli. Although these outbreaks generate a great deal of attention, probably many more fresh food contamination instances go undetected. While the agricultural supply chain is complex, new technologies are available that hold the promise of better track and trace for fresh foods. This article examines the trends and applications of RFID and related information technologies as a means of improving the safety for fresh foods in addition to being a platform for other innovations. Overall, data and information technologies will gain increasing importance in agriculture, affecting diverse areas such as food safety, supply chain, and marketing.

INTRODUCTION

In the 1967 the Surgeon General of the United States was quoted as saying, "The time has come to close the book on infectious disease… We have basically wiped out infection in the United States (Huber, 2007)." When the Surgeon General made this statement, antibiotics had substantially reduced the spread common infections. However, as bacterial resistance increased over time, it became clear that infectious diseases would continue as a threat to public health. In addition, fears began to arise that various pathogens responsible for causing food borne illness and other types of sickness might develop complete immunity to antibiotic treatments. The prospect of this development has heightened the need for various methods, techniques, and systems to monitor infectious disease outbreaks and to develop a means for responding. This is an ongoing priority for both governmental and business organizations as "Germs are always future, always reinventing themselves in their ingeniously stupid and methodically random way (Huber, 2007)."

For fruits and vegetables that are the un-processed product of farms, the issue of food safety relates directly to the movement of these products through the supply chain to the consumer. Defined as activities that include transportation, warehousing, purchasing, inventory management, and customer service, the supply chain takes on great importance in American business for both consumer goods and industrial products. The network of retailers, distributors, warehouses, and food processors that comprise the supply chain for fresh fruits and vegetables is effective in achieving fast, efficient delivery to consumers and represents a complex web of business activities. For example, spinach from the Salina Valley can be in 40 states within days of picking. This means that if an outbreak of food borne illness does occur, significant numbers of people become sick before governmental authorities can identify the cause (Quaid, 2006).

With the speed of modern supply chains, it becomes important to accurately track and trace products in the case of a recall because of defects causing immediate danger to consumers, inadvertent or intentional contamination, or the outbreak of food borne illness. *Tracking* involves knowing the real-time physical location of a particular item and its future path within the supply chain. *Tracing* is the ability to know the historical locations, the time spent at each location, record of ownership, packaging configurations, and environmental storage conditions for a particular item (Schuster et al. 2007).

To achieve track and trace for fruit and vegetable supply chains, two important components must be in place. First, there needs to be some sort of unique identification for the product. The means of identification must be fast enough for query during emergencies. Bar codes on packaging are a common means of identification within food supply chains; however, it is often the case that paper records are the primary means of identification. Many have advocated Radio Frequency Identification (RFID) as a way of achieving more accurate product identification. The technology is new and has some applications within the food industry.

The second requirement is interoperability of data between the different parties within the supply chain. This is a monumental task as data formats used by various computing systems differ and are seldom interoperable.

At MIT, significant previous and ongoing research has contributed findings and potential technological solutions to the problem of track and trace for the fresh food supply chain. A large amount of work has taken place as part of the Auto-ID Center effort (see autoid.mit.edu). This has included basic track and trace technology involving RFID and the application of sensors to monitor temperature while food products move through the supply chain. Some of these applications involved the U.S. military (Schuster et al. 2007).

Besides RFID, a five-year effort called the Data Center Program (see datacenter.mit.edu) seeks to develop a way to interoperate data and mathematical models. This research has culminated in the M Language, a useful approach to integrating large amounts of data in different formats, as is the case for fruit and vegetable supply chains (see mlanguage.mit.edu). The M Language achieves semantic and syntactic interoperability for data held in XML, a standard developed by the World Wide Web Consortium (W3C).

A new research group, the MIT Field Intelligence Lab (see fil.mit.edu), has the general goal of examining practical spatial problems experienced in industry. This effort includes a group devoted to agricultural systems productivity in addition to other subgroups that focus on environmental sampling, marketing spatial diffusion, and the study of the Megalopolis in Asia. A direct result of the agricultural research being conducted by FIL is a new way of spatial sampling and optimal sensor network design. This has applications in sampling fields of vegetables, vineyards, or orchards for potential pathogens that could pose a public health threat.

This paper consists of four parts. First is an overview of the elements of traceability, followed by a brief case-study discussion. Next is a description of RFID and its role in food traceability. The final section is an overview of data interoperability within the food supply chain.

THE ELEMENTS OF TRACEABILITY

Though food technology has advanced a great deal during the past 200 years, there continues to be concern about the safety of fresh foods that are the unprocessed product of farms. Prone to contamination E. Coli, and other bacteria, the freshness of fruits and vegetables is especially important in controlling outbreaks of food borne illnesses (Banwart, 1979). By one account, there are 76 million illnesses and 5,000 deaths each year from food contaminated with various pathogens (Thomsen and McKenzie, 2001). Because of several recent outbreaks involving packaged vegetables, there has been a growing interest in implementing tracing systems for fruits and vegetables. The basis of any system of this type depends on some form of unique identification.

For example, all agricultural supply chains share a commodity orientation where production of like goods takes place on numerous farms. The complexity of the agricultural supply chain arises because agricultural products often have significant variations in taste, vitamin, mineral, and protein content, bacterial contamination, and numerous other attributes that define the quality, safety, and identity of food. These variations sometimes depend on the location where production occurs, but it is also true that variation is a function of methods employed in agriculture, which do not always follow a standard practice.

Along with biological variation, there is also the inherent characteristic of raw material mixing. The output from farms must undergo various forms of processing before eventual consumption. This makes the task of maintaining unique identification throughout the supply chain difficult.

In the agribusiness environment, food traceability serves several different functions for various constituencies including identification of the origin of contaminated food (public safety), the limitation of liability in the event of disease outbreak (business), and information about inferred physical quality characteristics (consumer) (Hobbs, 2004). Most of the emphasis concerning traceability relates to safety and liability issues. Consumers tend to depend on branding for making quality and purchase decisions rather than a comprehensive understanding of the origin of specific food products. However, with an increased interest in local food production, this consumer trend is changing somewhat.

In addition to the underlying complexity of agricultural supply chains, there is also the established trend of globalization that adds new dimensions to traceability. Though there are many benefits from world trade, the crossing of borders increases the risks that various types of contamination might spread quickly worldwide (Unnevehr, 2004). Trading partners from around the world are increasingly interested in establishing agricultural tracing systems as a means of mitigating the negative economic consequences of disease outbreaks.

Perhaps the best way to understand the issues involved in track and trace is to examine a case study. The next section explores a food recall situation that occurred in 2003 and stimulated interest in developing a track and trace system. Though not directly related to the fruit and vegetable industry, this case highlights some of the difficulties in the rapid recall of food products.

A BRIEF CASE STUDY

In December 2003, a single case of Mad Cow disease was discovered in Washington State (Kilman et al. 2003). Almost overnight, more than 40 countries slapped a ban on beef imports from the US. On an annualized basis, beef exports equal $3 billion in revenue and about 10% of the domestic beef market (Stecklow and Kilman, 2004). Virtually all of the US export market was lost.

Upon discovery of Mad Cow disease in Washington State, US Department of Agriculture scrambled to find information about the extent of the problem. With no national system of cattle identification or tracking, the USDA sifted through paper records kept at local farms for information on the details surrounding the movements of the infected cow. The process required several weeks before an understanding of the magnitude of the problem became clear.

Initially, government investigators believed the infected Washington cow was among 80 that were raised on the same farm in Canada and eventually sold to various interests in the US. Since contaminated feed was probably the cause of the outbreak, USDA officials felt it was reasonable to conclude that the other cows in the herd were also exposed to Mad Cow disease and perhaps infected. At the conclusion of the investigation in February 2004, only 27 of the 80 were positively identified (Kilman, 2004). Likely, the other cattle in the herd had already entered the US food supply by the time the first case appeared. The USDA stated that it was impossible to determine which retailers might have sold the meat. In retrospect, a lack of information prevented development of timely answers to basic questions about the infected cow including; "Where did this animal come from, where were the feed sources, where did animals move out of this herd (Wald 2003)?"

The reality of modern agriculture is that feedlots, dairy farms, and other animal growing operations are potential breeding grounds for a number of diseases that could threaten the US food supply. This is also true for the fruit and vegetable industry where soils contain various pathogens that cause illness.

It is difficult to create a credible system for track and trace without some form of unique identification. What follows is an overview of identification systems and the need for unique identification as a base for a food track and trace system.

IDENTIFICATION TECHNOLOGIES

Considered one of the greatest innovations of the 20th century (Haberman, 2001), the bar code represents the first large-scale, automated effort to identify objects. Given these early successes, manufacturers began to adopt the bar code as a means of improving inventory accuracy and to coordinate supply chain operations. Product level data on time and place obtained from bar code systems greatly improved temporal and spatial utility within supply chains (Coyle et al. 1992; Simchi-Levi et al. 2002).

Since bar codes are now a mature technology, it is natural to look forward to the next stage in the commercial use of identification systems. The historical focus of bar codes has concentrated on identification of an object type. For a consumer goods item, this means the brand and size of an individual product, or the brand, size, and quantity contained in cases shipped from manufacturers to retailers.

Product type data, obtained with the ease of a laser scanner, provides enough information to automate checkout lines or to improve inventory management. The Universal Product Code (UPC), established by the Uniform Code Council (UCC) in 1973, serves an important role in establishing uniformity and order concerning the product type data read from bar codes (Haberman, 2001). As of 2005, over 1,000,000 organizations use the EAN/UPC on their products. This has unlocked enormous amounts of data to retailers and firms in other industries.

With an increase in the intricacy and sophistication of products, the needs of business are moving beyond identification of product type to unique identification of individual objects by serial number. This represents a significant transition because unique identification introduces a much greater degree of complexity in system management.

Yet at the same time, unique identification also offers a number of possibilities that include greater visibility and real-time control of objects located anywhere between the manufacturer and the customer. The full realization of these capabilities will most certainly revolutionize the practice of supply chain management.

Organizing for Unique Identification

The EPCglobal Network (see www.epcglobalinc.org) is a system operated by a nonprofit organization that builds on the tradition of automatic identification first established by the bar code. At the core of this system is the Electronic Product Code (EPC^{TM}), a serial numbering system designed to handle unique identification of trillions of objects. This numbering system forms a standard, uniform basis for linking physical objects together within a network that applies to all levels of the supply chain and across industries.

The means of creating such a network involves the placement of low cost RFID tags on objects such as cases, pallets, or individual products. In addition, RFID technology provides the capability for these tags to communicate with the Internet, secure Intranets, or point-to-point communication between organizations, through readers situated at various points within the supply chain. In the future, this type of network will form the base for "pervasive computing capabilities embedded in our everyday environments (Fensel et al. 2003)."

Creating a Global Standard

An important aspect of the EPCglobal Network is the development of common standards. Considered the bedrock of commerce, standards enable interoperability and the free flow of various business transactions. Often following technological breakthroughs, standards setting efforts have a positive record in driving economic growth that dates to the origins of trade in the ancient world.

During modern times, the co-development of new technologies along with the establishment of common standards represents a complex activity involving many different groups that must depend on each other for mutual success. Given expanding trade between countries, the interaction between new technologies and standards setting bodies will play an increasingly critical role in guiding the direction of innovation and future economic development.

As agribusiness and US government regulation agencies begin the process of discerning the importance of the EPCglobal Network to the track and trace problem, it is essential to know the fundamental aspects of the technology and to be able to abstract these capabilities to the practical applications of the present and future.

The Basic Elements of Unique Identification

The original designers of the EPCglobal Network had a predetermined idea of how to use the Internet as a means of implementing unique identification (Sarma, 2000). Early researchers also introduced the goal of interoperability across all levels of the supply chain.

The basic design called for three components. First, and perhaps most important, a low cost RFID tag serves as the base of the system. To reduce the cost of the tag, researchers focused on new methods that would turn the manufacture of tags into a mass production process.

A second major development involved placing a unique serial number on the RFID tag capable of identifying trillions of objects. This task involved the calculation of the size of the number to ensure adequate coverage. The numbering system developed during this phase of research became the EPC.

The last step was to build a computer infrastructure capable of processing identification data and information. Since initial designs of RFID tags called for minimum functionality as a means of reducing costs, the EPC became the only data contained on tags. It served as a pointer to greater amounts of data and information held in a network. In this way, an efficient balance was struck between the limited functionality (and cost) of RFID tags, and the ability of computer networks to hold important data and information about objects.

With these components in place - low cost tags, the EPC, and a network infrastructure for handling data – the EPCglobal Network is capable of new types of data and information exchange as compared to current bar code systems. In business, it is a general rule that the value of information increases when it moves beyond the four walls of a firm. Many firms from a wide range of industries are convinced collaboration provides extensive benefits and are looking for new technologies to enhance its application with trading partners.

The greatest value of the EPCglobal Network may be the set of standards necessary for supply chain wide communication of unique identification data. In this way, collaboration between trading partners can take place on a more sophisticated level and it becomes theoretically possible to monitor and communicate with objects at any step of a supply chain. This type of collaboration and control has definite benefits that are quantifiable once the EPCglobal system is in place.

In relation to the fruit and vegetable industry, RFID can have other favorable impacts on shelf life through tracking and tracing capabilities. Tanimura & Antle, a produce packer from Salinas, California, uses the tracing capabilities of RFID as an integral part of logistical decision-making (Wasserman, 2006). When farm workers harvest heads of lettuce, the product is wrapped and placed in a re-usable plastic container that contains an RFID tag. The EPC helps to identify where the lettuce was grown and the time that the lettuce was picked.

Given that harvest operations take place over a wide area, a number of trucks converge at the central refrigerated warehouse nearly simultaneously. Since chilling of the lettuce is the first step in packing, a bottleneck sometimes develops. When the tagged containers of lettuce arrive at the refrigerated warehouse, each is scanned and the EPC information recorded in a database. This information is vital in helping warehouse managers prioritize which containers to place into refrigerated storage first. By giving top priority to refrigerating the lettuce with the longest time since picking, the shelf life of this product can be extended. In the case of lettuce, getting the product chilled even one hour quicker can make a significant difference in extending shelf life. This limited application shows the potential of unique identification in fruit and vegetable agriculture.

Any sort of RFID effort will generate large amounts of data. Connecting this data in an interoperable way remains a challenge. The final section of this paper examines issues relating to data within supply chains, along with a new approach for interoperability called the M Language.

OVERVIEW OF DATA INTEROPERABILITY

Among other things, the M Language created at the Laboratory for Manufacturing and Productivity – Data Center Program, MIT, addresses a common problem encountered in industry, namely the existence of hundreds of incompatible data formats. Though the Extensible Markup Language (XML) developed by the World Wide Web Consortium (W3C) provides some standardization for storing and transmitting data, the standard also allows wide discretion in establishing data syntax (the order) and semantics (the tag for each data element).

Using an unambiguous vocabulary along with rules for syntax, the M Language provides a basis for interoperating data across disparate data sources. This provides the ability to integrate data from wide range of supply chain activities, decreasing the cost of data management and increasing the speed of retrieval. Drawing from linguistics, the M Language is similar to a synthetic language that exists between other established languages. The current version of M (at mlanguage.mit.edu) contains over 200,000 unique terms and 880,000 semantic relationships. Figure 1 shows a simplification of how the M Language integrates data across disparate sources.

In broad terms, the M Language forms a common link between distributed, incompatible data formats. With current information technology approaches, a programmer must first identify a target, and then establish a common semantic and syntax before communication can take place. Given all the different formats available, this task is time consuming and sometimes overwhelming. Often called point-to-point translating, the work of constantly comparing data semantics and syntax will eventually become unmanageable as the amount of data available increases and need for data sharing across the Internet expands.

The advantage of the M Language is that a translation needs to take place only once, at the edge of the computing system. Once in M, data exchange can take place with any other data that is also in the M format. In this way, a programmer can send data to an unknown target as long as it is in the M Language. Programmers do not need to know the format of every possible target for data exchange, achieving a new level of system integration.

This aids in the sharing of data within supply chains, which is a prerequisite for implementing track and trace technology.

CONCLUSION

The track and trace problem for the fruit and vegetable supply chain is an enormously complex problem that depends in unique identification and new ways of handling the large amounts of data anticipated from RFID or other types of identification technologies. Most of the information and electrical technologies needed to build robust applications of track and trace are still in an advanced form of research and development although some applications are now appearing in the fruit and vegetable industries. The prospect of an open system for the rapid means of identifying information related to a product recall is still several years into the future. Private systems currently exist that aid in track and trace, however, these are proprietary, and the information is not open. Industry is moving forward with software solutions but is reluctant to form a truly open system across all parties in the fruit and vegetable industry.

The basic technologies discussed in this paper, RFID, the EPCglobal Network, and the M Language are also the basis for the integration of marketing science, engineering technology, and supply chain management. This represents the next stage in delivering consumer products to customers though the supply chain.

Literature Cited

Banwart, G.J. 1979. Basic Food Microbiology. AVI Publishing Company.Westport, Connecticut.

Coyle, J.J., E. J. Bardi and C.J. Langley 1992. The Management of Business Logistics. West Publishing Company. New York

Fensel, D., J. Hendler, H. Lieberman, and W. Wahlster Editors 2003. Spinning the Semantic Web. MIT Press. Cambridge, MA.

Haberman, A.L., Editor 2001. Twenty-Five Years Behind Bars. Harvard University Press. Cambridge, MA.

Huber, P. 2007. The Coming Plague. The Wall Street Journal. April 10.

Hobbs, J.E. 2004. Information Asymmetry and the Role of Traceability Systems. Agribusiness 20:4:347 – 415.

Kilman, S, S. Leung, and T. Carlisle 2003. Case of Mad Cow Found in the U.S. for First Time. The Wall Street Journal. December 24.

Kilman, S. 2004. U.S. Concludes Mad-Cow Investigation. The Wall Street Journal. February 10.

Quaid, L. 2006. E. Coli Threats Show Food Chain Vulnerability. The Associated Press. Oct. 10.

Sarma, S., D.L. Brock, and K. Ashton 2000. The Networked Physical World. MIT Auto-ID Center.

Schuster, E.W., S.J. Allen, and D.L. Brock 2007. Global RFID: The Value of the EPCglobal Network for Supply Chain Management. Springer, New York and Berlin.

Simchi-Levi, D., P. Kaminsky and E. Simchi-Levi 2002. Designing and Managing the Supply Chain. McGraw-Hill/Irwin:New York.

Stecklow, S. and S. Kilman 2004. As U.S. Pleads Mad-Cow Case, Past Practices Are a Handicap. The Wall Street Journal. January 8.

Thomsen, M.R. and A.M. Mckenzie 2001. Market Incentives for Safe Foods: An Examination of Shareholder Losses from Meat and Poultry Recalls. American Journal of Agricultural Economics 82:3:526 – 538.

Unnevehr, L. 2004. Mad Cows and Bt Potatoes: Global Public Goods in the Food System. American Journal of Agricultural Economics 86:5:1159 – 1166.

Wald, M. 2003. U.S. Scours Files to Trace Source of Mad Cow Case. The Wall Street Journal. December 25.

Wasserman, E. 2006. Keeping Fresh Foods Fresh. RFID Journal 3:1.

Figure

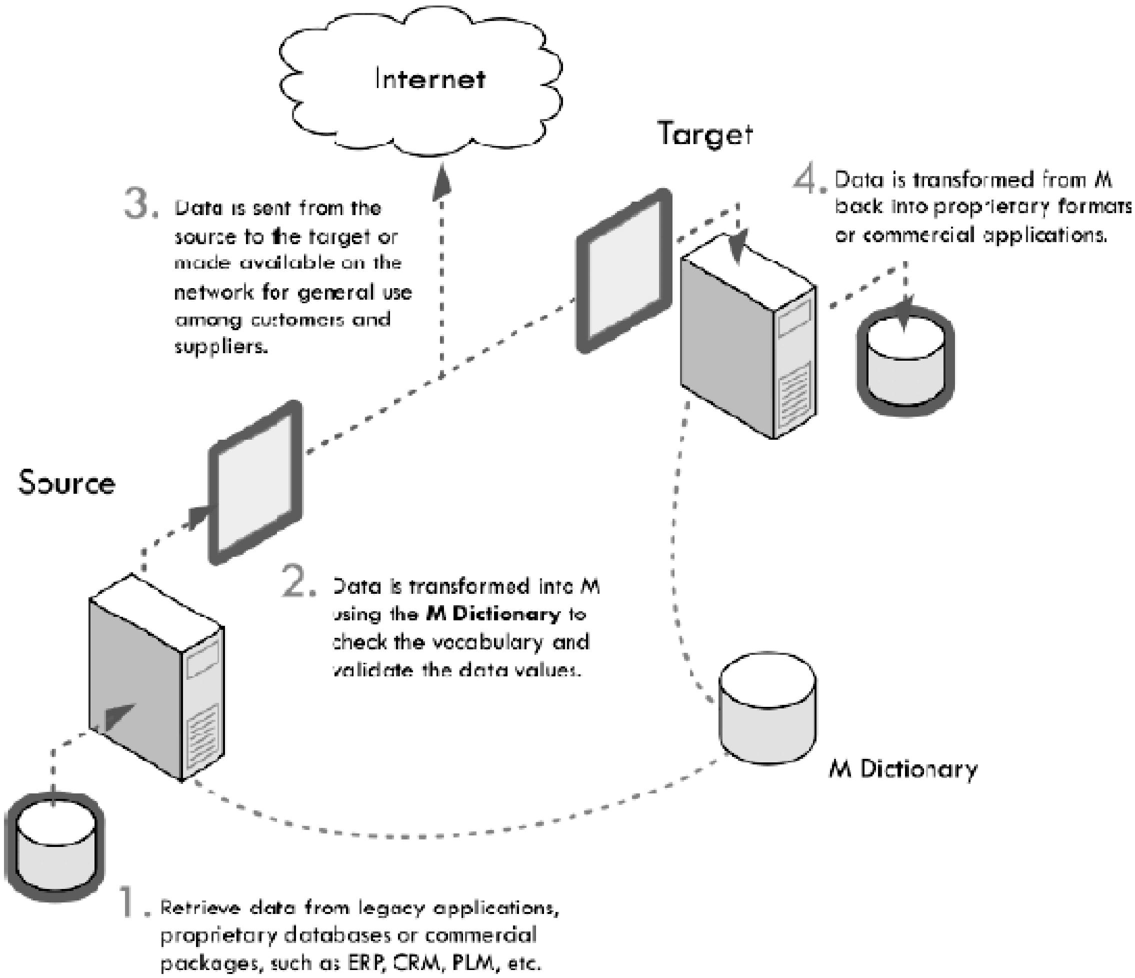

Figure 1. The M Language serves as an intermediary between disparate, proprietary data systems, as well as a general interface for internet communication.