CONSTRUCTING, VISUALIZING, AND ANALYZING A DIGITAL FOOTPRINT*

STEPHEN D. WEAVER and MARK GAHEGAN

ABSTRACT. Herein, we discuss the desire for new technology, the need for security, and the right to privacy; in doing so, we argue that each of these concerns comprises an important, tripartite debate. To highlight the complexities in this problem, we define our notion of a "digital footprint" and introduce BigFoot—specialized software created for the research described here to facilitate visualization and exploration of the data that comprise Stephen Weaver's personal digital footprint. Using BigFoot we demonstrate how multiple digital personae can be created from the data that constitute one unique digital footprint and provide a methodology for understanding the good and bad impacts that new technologies may have on future societies. One of the primary arguments of this work is that the debate—though not formally recognized—is currently before contemporary society and must receive sufficient attention. *Keywords: digital footprint, digital persona, privacy, security, surveillance.*

> Whenever a new technology is born, few see its ultimate place in society.
>
> —Paul Ceruzzi, [1986] 2000

In the age of pervasive, ubiquitous computing, concerns written about by science-fiction authors such as Aldous Huxley and George Orwell in the early twentieth century have found a new and poignant relevance, one that has escalated dramatically as events related to the United States' "war on terrorism" have progressed. Although the need for an improved intelligence infrastructure in the United States is undeniable, present threats to personal privacy have no historical analogue (except in science fiction). Concurrently, exciting new technologies that make use of federated databases and location-aware technologies are becoming increasingly available for all manner of uses. These technologies—as with those that comprise an effective intelligence infrastructure—have vast potential to both benefit society and erode personal privacy. It is imperative, therefore, to foster an informed public debate concerning the desire for new technology, the need for security, and the right to personal privacy.

In this article we discuss these three intermingled concerns; and in doing so we make an effective argument for the necessity of recognizing that all three are not only salient but also mandatory in discussions concerning the use and misuse of new location-aware technologies. Additionally, we define our notion of a "digital footprint"—the digital traces each one of us leaves behind as we conduct our lives—and describe our attempts to visualize this footprint using specialized software developed as part of this research. This article contributes to the privacy/security/technology debate by offering practical examples of the possibilities for constructing digital "profiles" of individuals and groups using current and emerging technologies. It

* This article grew out of Stephen Weaver's master's thesis, portions of which are reproduced with the author's permission.

MR. WEAVER is a doctoral candidate in geography at The Pennsylvania State University, University Park, Pennsylvania 16802, where DR. GAHEGAN is a professor of geography.

The Geographical Review 97 (3): 324–350, July 2007

seeks to provide a methodology for understanding the impacts, both positive and negative, that these new technologies will have on societies of the future.

We begin from the standpoint that technology of itself is inherently neither good nor bad. Nevertheless, whether a new technology should be put to a particular use can raise some extremely difficult questions. Implications following the introduction of some technologies can include ethical, legal, and moral issues that demand and deserve serious attention by all those who may feel—directly or indirectly—their potential impacts. It is our opinion that the continued use of existing surveillance technologies and the introduction of new technologies certainly warrant such attention. One of the contributions of the work described here is the method used to argue the importance of considering these issues as they pertain to technologies of surveillance, especially location-aware technologies of surveillance and the means to fuse together personal data collected across different organizations. It is our belief that the technologies discussed below have great potential for both harm and good. Furthermore, it is our belief that a holistic investigation of the potential impacts of these technologies is essential to ensure that human rights are not unnecessarily or unjustly violated in the pursuit of capitalist enterprise or national security.

A great many laws have been passed to ensure an individual's right to privacy; for example, the Fourth Amendment to the U.S. Constitution and the Data Protection Act of 1988 in the United Kingdom. However, since their inception, the rights that they seek to uphold have been challenged time and again. Challenges have come from the adoption of new technologies as well as the introduction of new legislation, such as the USA PATRIOT Act of 2001 (P.L. 107-56, [http://fl1.findlaw.com/news.findlaw.com/cnn/docs/terrorism/hr3162.pdf]). However, as we suggested above, these technologies also have many necessary and legitimate uses in both public and private sectors. The "war on terrorism," for example, has made the necessity of surveillance technologies unmistakably evident. Additionally, locational and mapping technologies such as GIS, GPS, and location-based services (LBS) have—despite some earlier objections (Pickles 1995)—become mainstream tools in many science, business, and government enterprises. Via cell phones, PDAs, and other fashionable gadgets, these technologies are becoming widely available to individuals. Coupled with the increased use of digital transactions and capabilities to integrate and mine the vast databases these transactions produce, it becomes increasingly possible to know not only where we are but also what we are doing—and with whom we are doing it. Thus the primary argument in this article is that a tripartite debate is currently confronting society. Can an appropriate balance be achieved between the desire for new technology, the need for security, and the right to privacy?

Technological Background

In writing about his bold vision of "ubiquitous computing" Mark Weiser claimed that "the most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" (1991, 1). One set of technologies that is beginning to weave itself into life's everyday fab-

ric—with both good and bad potential consequences—can be broadly labeled "GIS." During its infancy GIS was relegated to a few specialists, who used the software and hardware for a narrow set of applications, such as mapmaking and spatial analysis. Now, applications derivative of GIS are finding themselves in all types of both public and private spaces. In the last decade this has become increasingly true with the coupling of GIS with GPS, technology developed by the U.S. military for purposes quite different from those who exploit it in the private sector. For example, many private-sector uses of LBS are becoming increasingly popular—and profitable. Current uses of LBS technology include in-car navigation systems, such as OnStar, which provides, among other services, various forms of emergency assistance, remote diagnostics, stolen-vehicle tracking, driving directions, and, for a small additional fee, a personal concierge that can book hotels or suggest entertainment destinations (OnStar 2007). Another use of LBS technology includes mobile augmented reality systems, such as navigation systems for the visually impaired (Höllerer and Feiner 2004). These systems are designed to be self-contained and portable—that is, wearable—and help visually impaired individuals to navigate safely through familiar and unfamiliar settings without the assistance of additional guides.

Additionally, new databases and associated technologies are facilitating the cross-linking of massive personal data sets, with many benefits for individuals and industry. For example, businesses can target specific customers who are most likely to purchase their products or use their services. Geodemographics—the relationships between where people live and their likely demographic characteristics—enable marketers to predict behavioral responses of consumers based on statistical models of identity and residential location (Goss 1995). By marketing only to those customers most likely to purchase their products, businesses can reduce operating costs and thus offer lower prices.[1] In addition, businesses can access credit-rating scores to make informed decisions concerning whether to offer credit and, if so, how much. In the past, such decisions took days or even weeks to complete, but now, by accessing databases compiled from personal credit-card transactions integrated over time, businesses can make decisions in seconds (McCullagh 2004). Consumer profiles also allow credit-card companies to monitor accounts for suspicious activity and help prevent identity theft. Thus, in a market-driven society, both businesses and consumers benefit from the compilation and sharing of personal information; successful applications of the "new magic" of geodemographics are well documented (Robbin 1980; see also, for example, Hughes 1991; Thomas and Kirchner 1991; Baker and Baker 1993; Curry 1993).

Security benefits also result from the combination of massive, cross-linked data sets and location-aware technologies. For example, one could argue that this combination provides security against terrorism and can help prevent unlawful acts that might endanger the public welfare. In the United States, the Department of Homeland Security, with the vision of "preserving our freedoms, protecting America . . . and securing our homeland" (DHS 2004, 4), and the multitude of contractors supporting that vision put significant pressure on both the supply side and the de-

mand side of these technologies. However, as much of society accepts these technologies, astoundingly little debate has emerged in society at large regarding social implications—with notable exceptions in the academic literature (for example, Dobson 1998, 2000, 2002; Lyon 2001; Dobson and Fisher 2003; Farmer and Mann 2003a, 2003b)—and in 2002 the University Consortium for Geographic Information Science declared the social implications of LBS a short-term research priority (Kim 2002).

With Great Britain's nearly 2 million closed-circuit television cameras installed in public places as one obvious example (EPIC and PI 2002), surveillance is rapidly becoming commonplace and thought of as just another utility—like water or electricity. Individuals and organizations on the demand side of this burgeoning market span the spectrum from the private and public sectors, including concerned parents who want LBS to track the whereabouts of their children, "patriotic" citizens who dare not criticize technology that could potentially prevent terrorist acts; eager governments that wish to exercise control over their citizens and interests, opportunist contractors who wish to cash in on lucrative security contracts, enterprising businesspeople who wish to know as much about their customers as possible in order to minimize costs and maximize profits, and "average" individuals who enjoy the benefits of new technology without giving much consideration of or appreciation for the social and ethical tradeoffs resulting from supporting its development and deployment.

Put succinctly, these technologies are being put to market primarily for five reasons: people want them; governments demand them; suppliers profit from them; debates concerning their development and laws concerning their deployment cannot keep up with their advancement; and people's reasonable expectation of privacy is diminishing. However, as Jerome Dobson and Peter Fisher caution, the "countless benefits of LBS are countered by social hazards unparalleled in human history" (2003, 1). Furthermore, they define "geoslavery" as "a practice in which one entity, the master, coercively or surreptitiously monitors and exerts control over the physical location of another individual, the slave" (pp. 1–2). As futuristic and fantastical as this idea may seem, geoslavery finds its roots even deeper than the unnerving ideas conjured by Orwell and his contemporaries (1949). The technologies that provide navigational assistance to the visually impaired facilitate improved search-and-rescue capabilities, enable smart-shelf technology, and lead to lower prices for consumers are also those that are needed to monitor, record, and, to a certain extent, predict and control human behavior and location. Many of these technologies are now fully developed and commercially available. As a result, "geoslavery now looms as a real, immediate, and global threat" (Dobson and Fisher 2003, 1).

The French philosopher and social critic Michael Foucault asserted that isolation and the mere suggestion of being watched are all that is necessary to exert control over a physical body. He also made wrote about the panoptic power over the mind; for example, "Without any physical instrument other than architecture

and geometry, [the Panopticon] acts directly on individuals; it gives power of mind over mind" ([1975] 1995, 206). In modern society our spaces are organized "like so many cages, so many small theatres, in which each actor is alone, perfectly individualized and constantly visible" (p. 200). The asymmetry of spaces with unobservable observers was the very essence of power for Foucault because, "ultimately, the power to dominate rests on the differential possession of knowledge" (1983, 223).

Foucault's description has led to many treatises focusing on the centrality of vision in his metaphor. We need not, however, worry about either Orwell's or Foucault's conceptions of Big Brother; rather, we need to be concerned about the wide-reaching network of little brothers, sisters, and cousins—the many private corporations and businesses that, at least in the United States, often operate with fewer privacy controls vis-à-vis central government.

Who gains access to the information trails we leave behind and how widely the contents are shared, traded, and used is often unclear. With technologies such as LBS, user privacy could be carefully engineered into the service architecture, thereby reducing opportunities for misuse. Moreover, it is also possible to allow users to customize their privacy policies depending on contextual situations, for example, so that privacy policies do not degrade LBS capabilities.[2]

The Electronic Privacy Information Center has collected data from a variety of public-opinion surveys conducted in the United States between the early 1990s and 2002 (EPIC 2005). The findings can be summarized thus:

- Individuals should be in control of both initial data collection and data sharing.
- Individuals want accountability and security.
- Individuals want comprehensive legislation, not self-regulation.
- Individuals value anonymity.
- Individuals do not trust companies to administer personal data and fear both private-sector and government abuses of privacy.
- Individuals are unaware of prevalent tracking methods.
- Individuals want notice of how their personal information is collected and used and with whom it is shared.

With respect to public opinion on location privacy, when survey respondents are initially questioned, they are very concerned with their privacy. However, when the same individuals are using LBS—and thus benefiting directly from data services—their concerns appear to diminish (Taylor 2003; Barkhuus 2004).

The Digital Footprint

Personal information has become increasingly digital; that is, it has been either digitized from analog sources or created and stored in digital format. The last several decades have seen a huge amount of work dedicated to linking such disparate information together, using federated database technology, geocoding, ontologies, and query rewriting (Sheth and Larson 1990; Guarino 1998; Kashyap and Sheth 1998;

Egenhofer 2002; Bench-Capon, Malcolm, and Shave 2003). In the discussion that follows we imagine the situation in which all of this information, every shred of digital data—every movement, every transaction, every record—is woven together into a single virtual database. For a given individual, we refer to his or her accumulated trace as a "digital footprint."

The social and ethical ramifications of creating such a database have long been discussed. For example, when the U.S. government proposed the creation of a National Data Center in the mid-1960s, Representative Frank Horton (Republican, New York) said: "One of the most practical of our present safeguards of privacy is the fragmented nature of present information. It is scattered in little bits and pieces across the geography and years of our life. Retrieval is impractical and often impossible. A central data bank removes completely this safeguard" (Rule and others 1980, 56).

Although Horton's statement remains true and relevant, much has changed since he had the floor of the U.S. House of Representatives. First, a "central data bank" is no longer necessary to remove the fragmentation safeguard. Information can now be stored in disparate databases and transparently joined locally or over the Internet (Ludaescher, Gupta, and Martone 2001; Klein 2002). Second, processor speed, storage capacity, and network-transmission capacity—in addition to the development of database-management systems and data-mining techniques—have enormously increased in sophistication and will continue to do so for many decades to come. This situation has led Dan Farmer and Charles Mann to claim that "by 2023 large organizations will be able to devote the equivalent of a contemporary PC to monitoring every single one of the 330 million people who will be living in the United States" (2003a, 38). Third, the amount and nature of networked, linked, and searchable personal data have vastly increased as a result of efforts by both the private sector and the public sector, coupled with the growth in electronic transactions.

Several attempts to list the categories of personal information linked in public and private databases can be found in the literature (see, for example, OTA 1987; Gandy 1993; Goss 1995). These lists, though quite detailed, are insufficient descriptions of a digital footprint because of the rapidity of technological advance and the potential for profit via the collection of personal data on individuals. Large private-sector companies, such as ChoicePoint ([www.choicepoint.com]) and LexisNexis ([www.lexisnexis.com]) have realized the value of personal data and have begun compiling vast amounts of them. These data are routinely sold to interested parties in the private and public sectors. Individuals can also gain access to personal data. For example—and for a fee—one of the growing number of online vendors that sell collections of individual personal data provides information on

> adoptions and birth parents, ancestry archives, arrest records, background checks, bankruptcy records, birth records, child-care and nanny screening, correctional files, court records, courthouses, credit reports, criminal files, criminal indictments, death records, district-court files, divorce records, Department of Motor Vehicles records, driving records, driving-under-the-influence [of alcohol or drugs] files, driving-while-intoxicated records, estate records, family-history detectives, Federal Bureau of In-

> vestigation files, federal-court dockets, fraud and alias files, genealogy detectives, identity-theft records, incarceration and felony arrests, inmate locators, judgment files, juvenile files, lien records, marriage records, military records, missing people, naturalization records, offender records, parole records, people searches, plaintiffs/defendants, police records, probate records, property records, public records, real-property ownership records, sentencing files, sex offenders, skip tracing, small-claims records, superior-court records, unlisted home and cell-phone numbers, vital records, warrant files, and more. ([www.identitycrawler.com/what_you_get.html])

Thus, although the contents of a digital footprint may be difficult to elucidate and constantly evolving, we can say with certainty that it is a high dimensional and constantly growing space characterized by digital transactions, augmented by surveillance, and influenced by associations and patterns through space and time.

Experiments with a Digital Footprint

Researchers have begun to address issues of utility and privacy of locational information by using personal GPS devices to create space-time trajectories for individuals gathered over days or weeks (Raper, Rhind, and Shepherd 1992; Kwan 2000; Mountain and Raper 2001; Mountain and Dykes 2002). For the research described here, the data-collection process was significantly more thorough and elaborate (and time consuming) than that in earlier studies. In addition to continuously logging position, we included all sources of digital information to which we had access—from government-held Social Security records, through credit-card transactions and credit-score reports, to uses of public transport, libraries, and medical services. This involved the identification and characterization of the myriad sources of personal digital information that may be available to interested parties (private corporations, private individuals, or government agencies, for instance), given a technological climate characterized by ubiquitous tracking and surveillance.

To create a digital footprint we collected all the data Weaver left behind while going about his normal daily life between 6 November 2004 and 4 February 2005. After we synthesized the data, we posed several research questions. How should such data be displayed or visualized? Which aspects of behavior and persona are well represented in the data, and which are not? Is it possible to create hypothetical digital personae and find them in the database? What kinds of errors and uncertainties in the data may cause incorrect interpretation of activities by a third party? We investigated these and other questions using a custom-built software application that facilitated visualization and analysis of the digital footprint, then developed several plausible activity scenarios to facilitate discussion of the implications of new dimensions of personal data—for example, location—being linked to traditional point-of-sale (POS) and demographic data.

The software, called "BigFoot"—an amalgam of "Big Brother" and "Footprint" (not to mention the creature that would be a great deal more frightening if its existence had been proved)—was designed to show how space-time trajectories and digital

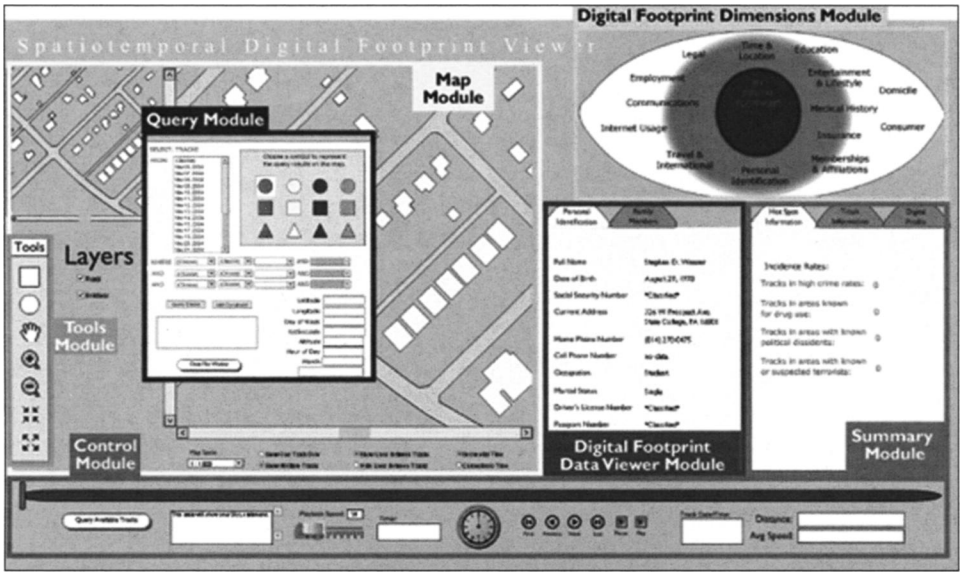

FIG. 1—As this screen capture shows, BigFoot consists of seven modules that facilitate the visualization, exploration, and analysis of a digital footprint.

transactions contribute to the creation of digital personae (Figure 1).[3] To that end, BigFoot was also designed to facilitate spatiotemporal queries about digital footprint data.[4]

TRACKING LOCATION

Weaver kept a Garmin Foretrex 201 GPS unit with him at all times to record his latitude, longitude, altitude, and time at 10-second intervals. The 67,000-plus track points collected during the ninety tracking days constitute the Time and Location dimension of the digital footprint (Figure 2).

Note that the GPS unit was used here as a surrogate for other position-tracking technologies, particularly cell phones. Cell phones can compute their position via GPS or cell-phone-tower triangulation. Hence, it is not necessary to plant some kind of tracking device on a subject; most of us carry them willingly! Other kinds of tracking technology that could be used to report position include surveillance cameras—such as those now commonly found on highways in the United States or the nearly 2 million cameras installed in public spaces in London, England (EPIC and PI 2002)—and the locations of POS transactions.

TRACKING ACTIVITIES

POS data are an essential component of a digital footprint. Many articles and books tout their utility for describing how individuals spend their time, their interests, what they consider important, and their opinions on various topics, both current and projected (Engel, Blackwell, and Kollat 1978; Michman 1991; Raper, Rhind, and Shepherd 1992; Curry 1993; Goss 1995). Data pertaining to Weaver's electronic POS

FIG. 2—The gray circles in this screen capture show approximately 28,900 of the 67,066 track points collected. Most of the white polygons indicate Pennsylvania State University buildings; Beaver Stadium can be seen in the upper right-hand portion of the image.

transactions collected during the tracking period included all transactions conducted with credit or debit cards, cash transactions with supermarkets, video renters, and other assorted retail businesses, "cashed" personal checks, transactions conducted at automated teller machines (ATMS), online payments, and electronic deposits and fund transfers to bank accounts. In addition, receipts were scanned and saved in a folder from which BigFoot could load the images as needed. This functionality was included for several reasons. Not only does it help demonstrate the potential invasiveness of systems such as BigFoot, but many companies are now saving receipt images in private databases, so providing such an image is realistic.

Other collection efforts included recording data from: transactions with libraries, film renters, and insurance companies; legal and medical histories; and profes-

TABLE I—PRIMARY DIMENSIONS AND EXAMPLES OF ASSOCIATED SUBSPACES OF THE DATA USED TO CREATE A DIGITAL FOOTPRINT

PRIMARY DIMENSIONS	EXAMPLES OF ASSOCIATED SUBSPACES
Affiliations	Political, religious, volunteer, and professional memberships and activities
Communications	E-mail address book; e-mail messages sent and received; telephone calls sent and received
Education	Institutions attended; areas of study; advisors and other associates; degrees conferred; academic transcripts; activities
Employment	Employment history; employee reports
Entertainment and lifestyle	Library records; film rentals; newspaper and periodical subscriptions
Finances	Internal Revenue Service records; bank accounts; property ownership; credit rating; automatic-teller-machine/credit-card/debit-card transactions; checks written and received; taxes; loyalty-card profiles
Insurance	Insurance policies
Internet usage	Internet service provider; Web sites visited; online accounts and profiles
Legal documents	Driving record; trafffic violations; criminal record; court records; attorneys' records
Medical documents	Medical history; prescription medications; illnesses; vaccinations; health-care providers
Personal data	Personal identification information; family members; Social Security and other identifying numbers; driver's-license number; passport number
Residence	Current and former places of residence; neighbors; associates; utilities usage
Time and location	Space-time trajectory
Travel and international data	Travel history; U.S. Immigration and Customs Enforcement records; international associates; international financial transactions

sional, religious, and volunteer affiliations. Table I is a summary of the data dimensions—and examples of associated data subspaces—that were used to create the digital footprint to be visualized in BigFoot.

BIGFOOT

BigFoot comprises seven primary components that work together in various ways to facilitate visualization and exploration of a digital footprint by performing a host of operations on external data that is loaded into BigFoot at runtime (see Figure 1).

1. The Query Module has several components (Figure 3). A series of drop-down boxes that build a query using structured query language as a model. After a query is built, sent, and received, the results are parsed into arrays that BigFoot

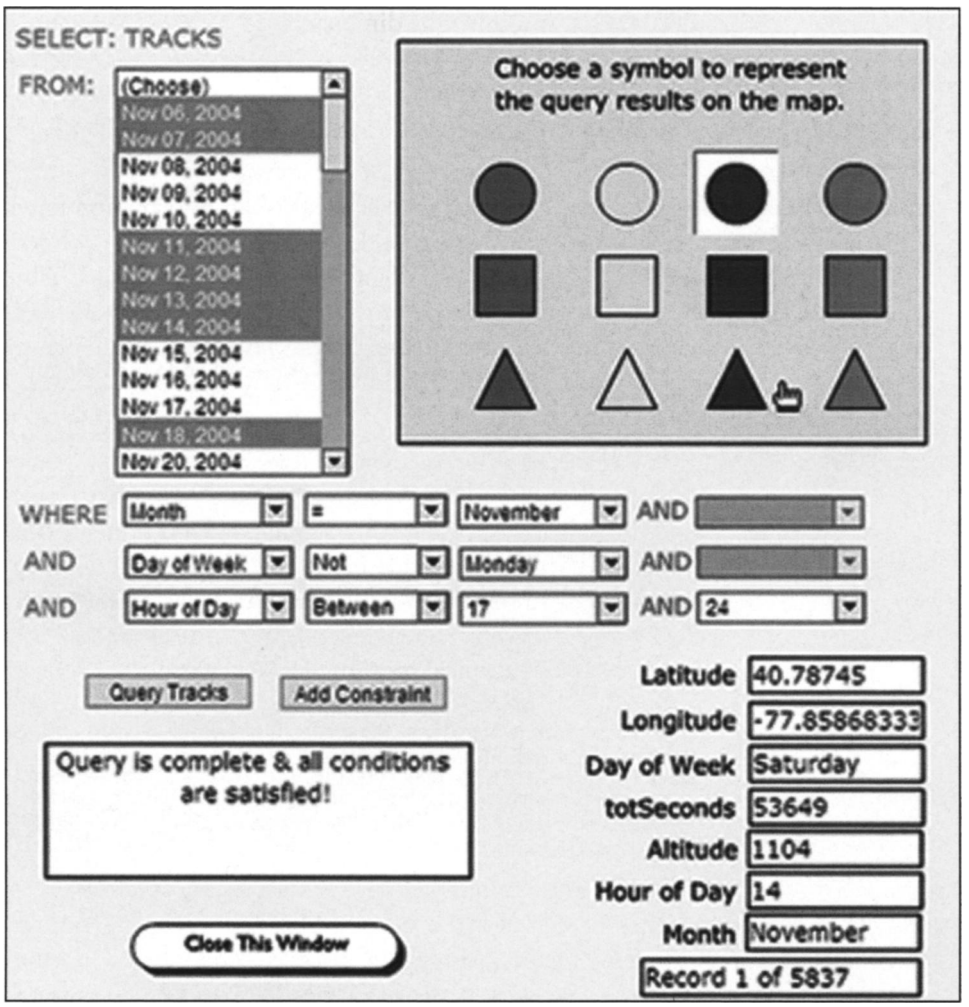

FIG. 3—The Query Module was designed to be robust yet user-friendly. It facilitates many kinds of temporal queries about the space-time trajectory data in the digital footprint.

can use to plot tracks in the Map Module (see item 6 below). The box containing point symbols (in the upper right quadrant of Figure 3) permits the operator to select a symbol to represent the query results in the Map Module. Because of this functionality, an operator can overlay the results of multiple queries and investigate the resulting spatial patterns within and between the query results. For example, an operator can overlay the tracks from consecutive Monday mornings to look for behavioral deviation or overlay the tracks from multiple subjects to look for interesting patterns or intersections among the subjects.

2. The Digital Footprint Dimensions Module is essentially an interactive diagram of the data dimensions that constitute the digital footprint. When an

operator moves the cursor over any of the dimension labels, the labels change color and the subdimensions that make up the digital footprint space (of the moused-over dimension) are shown in the middle of the diagram. If the user clicks on the dimension label, information is sent to the Digital Footprint Data Viewer Module.

3. The Digital Footprint Data Viewer Module consists of an information panel with dynamically configured tabs. Initially, the tabs correspond to the subdimensions of the Personal Data dimension of the digital footprint (see Table I). If an operator clicks on one of the dimension labels on the Digital Footprint Dimensions Module, the tabs of the Digital Footprint Data Viewer Module reconfigure themselves based on the information they receive from the click event. Like the information panels of the Personal Data dimension, appropriate data are retrieved from external sources, parsed, and displayed.
4. The Summary Module contains an information panel used to display interesting signals found in the data. An interesting signal can have many forms, depending on the data, the operator, and the operator's purpose. For example, an antiterrorism official may wish to flag certain books or authors as interesting signals. Other interesting signals may result from the space-time trajectory data. For example, an atypical movement pattern or frequent visits to hot-spot areas (see item 7 below) may be deemed interesting. Although the definition and retrieval of interesting signals is best when sophisticated algorithms are utilized, that method was beyond the scope of the research described here. Instead, several interesting items were hard coded into BigFoot to demonstrate how this functionality may work in later versions of the software.
5. The Control Module is located near the left-hand bottom of the BigFoot application (see Figure 1). As its name suggests, it contains buttons and other assorted widgets to control much of the interactivity between the operator and BigFoot; for example, some of the buttons control "playback" of the query results. In addition, dynamic text boxes display the date and time properties of the most recent track point mapped and the distance and average speed between the most recent track point mapped and its immediate predecessor.
6. The Map Module displays the results of the spatiotemporal queries created in the Query Module (see item 1 above).
7. The Tools Module comprises a "dragable" toolbox containing several buttons that permit the operator to manipulate the visualization. First, the summary tool allows the operator to draw a selection area (as a semitransparent rectangle) on the map in order to obtain summary information derived from the selected points. One potential benefit of this functionality is the ability to investigate a cluster of points quickly (Figure 4). For example, if BigFoot is currently displaying all track points that occur on Monday mornings and an interesting cluster is found that is not near the subject's residence or place of

FIG. 4—When the Summary Module is superimposed on the Map Module, several point clusters become visible. This screen capture illustrates the result—in the Summary Module—of using the summary tool to draw a box around points or, in this case, clusters of interest.

work, the BigFoot operator may wish to determine whether this activity is a recurring event. Rather than performing multiple queries, the operator can draw a selection rectangle that quickly summarizes the interesting cluster. Because other interesting artifacts may be found, the summary tool facilitates a kind of exploratory data analysis.

The hot-spot tool allows the operator to draw semitransparent circles on the map that will act like hot-spot areas. These hot spots can represent, for example, buffers around critical facilities such as nuclear plants or various types of zones, as defined by law-enforcement data spaces. For instance, an operator may have criminal-activity data and locations of suspected terrorist safe houses. After drawing hot spots in the appropriate locations (on the Map Module), the operator can perform

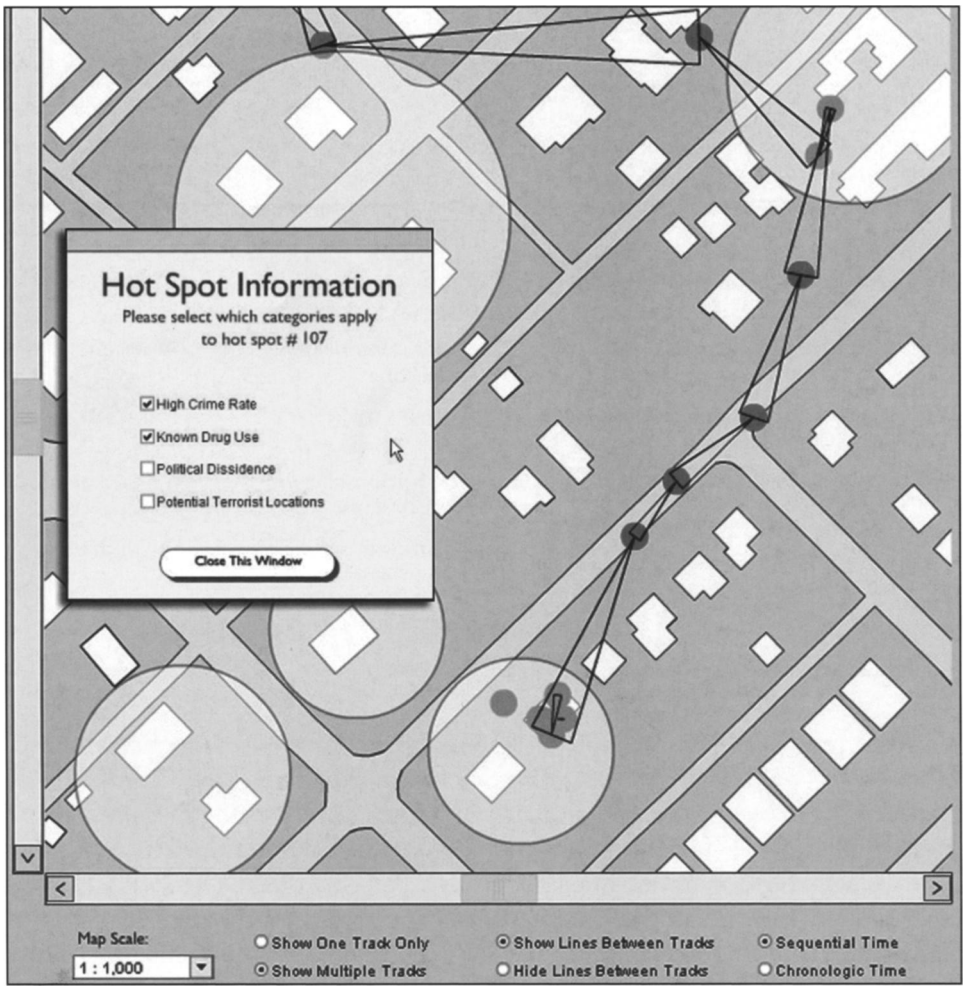

FIG. 5—An operator uses the hot-spot tool to determine whether the subject intersects locations known to have a high crime rate and a high degree of drug use. With the directional triangles shown between the tracks, the operator can see which hot spots were intersected, in what order, how often, and so forth.

spatiotemporal queries to quickly determine whether the subject encountered either or both of these areas (Figure 5).

TRACKING A SUBJECT'S MOVEMENTS AND ACTIVITIES

When an operator views the results of a query—that is, a space-time trajectory—the output is linked to the data that constitute the activity. As a result, the operator can concurrently view a space-time trajectory and any transactions conducted in that spatiotemporal vicinity. So, for example, if any type of transaction—a credit-card purchase or a library-book loan, say) is found during tracks playback, BigFoot will alert the operator and display details of the transaction(s).

TABLE II—SCENARIOS IN WHICH TIME AND LOCATION DATA RESULT IN ETHICAL ISSUES

SCENARIO[a]	EXAMPLES/REASONS	ETHICAL ISSUE
They don't know where you were		Absolute privacy
They know where you were, and you want them to know	On call, emergency, alibi	Consent
They don't know where you were, but you think they do know	Faulty signal, signal blocked, physical density	Ignorance
They know where you were, and we want them to know	Criminal activity, unethical conduct	Public interest
They know where you were, and you don't want them to know	Nuisance calls, stalking, monitoring	Invasion
They know where you were, and we think we want them to know	Abusers, hostage takers	Coercion
They know where you were, and then can physically keep you there	Geoslavery, forced labor, domestic abuse	Enslavement
They think they know where you were, but they are wrong	Surveillance, provision of location-based services	All of the above

Source: Adapted from Fisher and Dobson 2003.

[a] "They" refers to the people using the data; "you" refers to the person about whom data were collected and are being used; "we" refers to the wider society.

To demonstrate how BigFoot can be used to track a subject's movements and activities, we will begin by investigating the space-time trajectory recorded on 25 December 2004, which an operator may interpret as interesting because, on the majority of available tracking days, the subject did not travel outside Centre County, Pennsylvania, the county in which 51,727 (77.1 percent) of the 67,066 tracks were recorded. By playing back the tracks, the operator sees that the subject started in Gap, Pennsylvania, at 1:29:06 P.M. Then, after traveling for approximately 2 hours and 21 minutes, a brief stop was made in State College, Pennsylvania. Next, travel resumed for approximately 2 hours and 13 minutes, when the subject reached a destination in Oil City, Pennsylvania.

A SIMPLE JOURNEY HOME

Taking into account the facts that 25 December is a national holiday in the United States, that the subject was currently a student attending Pennsylvania State University (known from the data comprising the Education dimension, and that the subject was born in Oil City (known from the Residence dimension), the operator may infer that the subject was traveling to visit family members at Christmas. So why was the subject in Gap? The operator knows the details of other family members from the Personal Data dimension. Likewise, lists of associates are available via the Residence dimension.[5] After browsing the data from the Finances dimension that is in the same temporal vicinity as the tracks near Gap, the operator finds nothing out of the ordinary: an ATM withdrawal in the amount of $30.00; a small purchase at a bookstore; and a gasoline purchase at a Wawa Food Market.

After looking at the data in other dimensions of the digital footprint, the operator discovers that the subject has neither memberships in, or affiliations with, any organizations nor involvement in any legal disputes near Gap. However, after looking at the detailed data in the Communications dimension, the operator finds some useful information—several telephone calls to Lancaster, Pennsylvania, which is less than 20 miles from Gap. The operator may deem this interesting and direct further research to the recipient of the calls.

Although the scenario just elaborated is trivial, it demonstrates the analytical and exploratory capability of BigFoot. With BigFoot, an operator can test a series of hypotheses. The result of this testing is a series of clues that will often point to an answer or to specific avenues for additional research. Only by having immediate access to diverse data sets can some of these more difficult connections be made and questions be answered.

POTENTIAL PROBLEMS CAUSED BY POSITIONAL ERRORS

Although they focused solely on the Time and Location dimension, Fisher and Dobson have raised many serious concerns related to the use of surveillance data. In their 2003 article they describe seven scenarios in which data that make up the Time and Location dimension of a person's digital footprint are used and the ethical issues that result; Table II is our adaptation of their description, to which we added one scenario they did not consider: "They think they know where you were, but they are wrong. As shown below, an example may be due simply to locational error in the data.

A huge problem with surveillance based on hypothesis testing and profiling is that of balancing errors of commission against errors of omission. In this case, errors of commission refer to problems with the methodology, data, or interpretation that would cause a subject to be incorrectly associated with a particular hypothesis; for example, "X is a drug user." Conversely, errors of omission arise when a true hypothesis is not asserted. Errors of commission routinely occur when we use generalizations—such as profiles—to interpret individuals' actions. Some additional errors should be expected due to simple coincidence; and yet more errors, to inaccuracies in the data. Although all data used in these experiments are prone to errors and uncertainties, the two largest sources of error we encountered were the purchasable online profiles and the GPS data. The accuracy of GPS varies a great deal as a subject moves around a built environment because the infrastructure can obscure a varying subset of the satellites potentially in view. The effect is a constantly varying error component, and the result is that, when tracks are plotted on a map, they imply a subject went to—or past—places that in fact they—may have—avoided. For example, Figure 6 shows a subset of tracks recorded on 12 December 2004. Due to location error, several of the tracks make it appear as though the subject entered several buildings along the route, none of which was actually entered. Moreover, because the subject often walked to work, the recorded tracks shown in the figure are very close together. That increases the chance that multiple tracks were recorded

in or near a place where the subject did not go and would not want others to think that he or she had gone.

The consequences of inaccurate location data can be extremely damaging. The procedure for removing some of the extreme outliers from the tracking data is quite straightforward, but the likelihood of removing all inaccuracies from a data set created from the continuous tracking of an individual is unlikely unless much more adept tracking technology is used (for example, with a more sophisticated GPS receiver, an estimate of spatial accuracy would be derived at each tracking point). In the following scenario, we illustrate how a series of seemingly benign coincidences can lead to false interpretations of a digital persona.

A MORE SERIOUS SCENARIO

In this scenario a series of arrests for ownership of illegal firearms has caused the local police department to begin surveillance on a suspect's home. Coincidentally, the subject passes the suspect's home as he walks to and from work each day. This alone would not be cause for suspicion. But imagine also that a thief steals the subject's wallet and uses the cards contained therein to purchase books or to borrow volumes that have been flagged in a Federal Bureau of Investigation database (lending records from public libraries in the United States have been used in many FBI investigations [Foerstel 1991]). Moreover, a credit card is used to rent a van. Unaware of the loss, the subject's spouse goes to the local lawn-and-garden store to buy nitrogen-rich fertilizer for the family's garden. This small set of coincidences can be enough to cause an investigator to examine the subject's digital footprint in more detail. Because the subject's digital persona has already become suspicious, location errors and other possible coincidences found in the digital footprint may take on more significance than they merit.

In another scenario, a BigFoot operator wishes to build a profile of the subject's habits because of interesting correspondence found in the Communications dimension of the digital footprint. To begin, the operator defines three temporal ranges—morning, midday, and night—and commences analysis with the subject's morning activities. A query of all tracks that occur between the hours of 5:00 A.M. and 12:00 P.M. results in a set of 3,584 points. The operator then notices that 2,551 (71.2 percent) of the tracks are within the bounds of Centre County and 1,662 of these points (46.4 percent of the total) are in the Borough of State College. Of the 1,033 points outside Centre County, 477 are along a path that begins on Interstate 80 and terminates in Oil City. From the Personal Data dimension of the digital footprint, the operator knows that the subject was born in Oil City and, consequently, may infer that the subject traveled to visit relatives. The operator probably also infers that the subject traveled by automobile or bus, because the average speed for these tracks was 40.0 miles per hour and the median speed was 51.8 miles per hour.

All of the remaining points outside Centre County are in neighboring Huntingdon County and were recorded on Friday, 4 February 2005 between 10:09 A.M. and

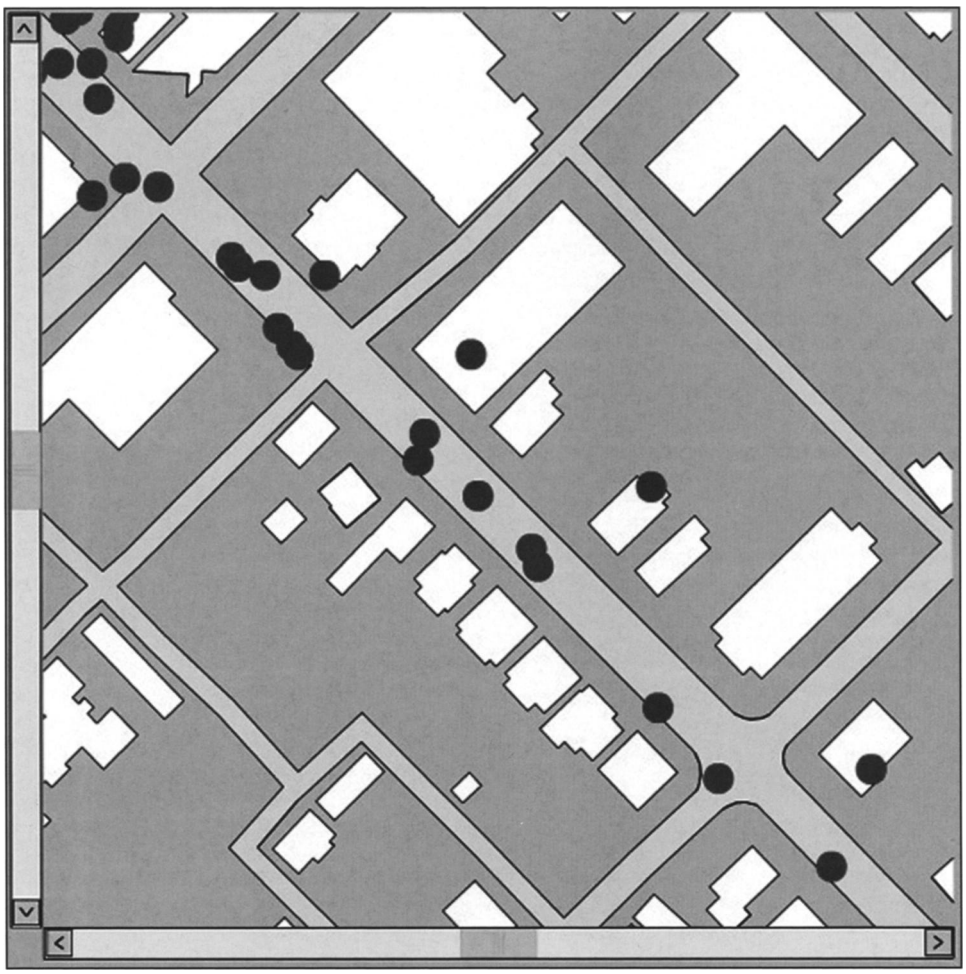

FIG. 6—Tracks recorded during a leisurely walk to the Pennsylvania State University campus on 12 December 2004 demonstrate how GPS errors can lead an operator to infer—incorrectly—that a subject entered one or more locations along a route.

11:45 A.M. These points may cause suspicion and would likely be marked for further investigation. However, before considering this anomalous activity, the operator may first want to determine what usual activity is for this time of day. To do that, attention is first directed at those points located in Centre County.

At first glance, the operator may notice several interesting things. The tracks heading south toward Huntingdon County are those recorded on 4 February 2005 (and noted above). However, one additional observation about these points is that a different route was taken on the return to Centre County. North of these points are three interesting paths: one to the northwest toward Park Forest Village; another to the northeast toward Houserville; and one to the east along State Highway 192. Because the operator knows from the Residence, Education, and Employment di-

FIG. 7—The points of this interesting path seem to indicate that the subject's destination was in and/or behind a building located northwest of the Borough of State College.

mensions that the subject lives, attends classes, and works in the Borough of State College, determination of usual morning activity will be focused on this region. However, the operator may first wish to conduct a cursory examination of the three new interesting paths to determine whether they warrant further investigation.

The tracks in the first path were recorded on 16 November 2004. Thus, between the hours of 5:00 A.M. and 12:00 P.M., this trajectory is unique. By animating the trajectory in BigFoot, the operator can see that the path was toward some destination northwest of State College and back, along the same route. Using the Tools Module to zoom in on the destination point, the operator can see several points located in front of and behind an unidentified building (Figure 7). The operator is careful to remember that accuracy and precision of the space-time data, as well as that of the other layers, warrant concern at this scale. However, the operator may

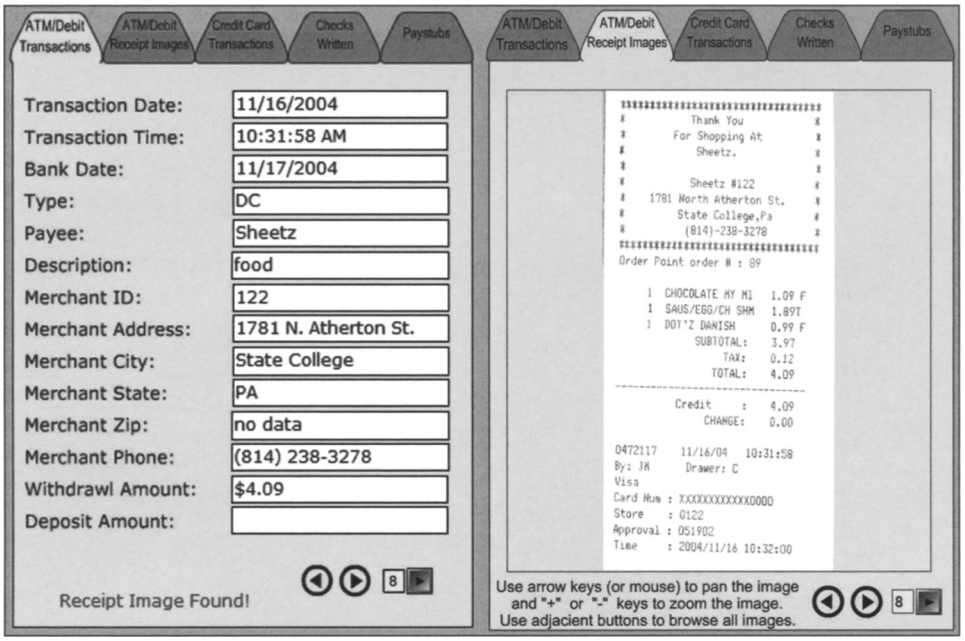

FIG. 8—In each of these two separate instances of the Digital Footprint Data Viewer Module the viewer is configured to show data comprising the "consumer" dimension of the subject's digital footprint. However, in the first instance (*left*) the viewer is showing transaction data in tabular form; in the second instance (*right*) the viewer shows the same data as an image of the receipt generated during the transaction.

begin to wonder whether the subject met someone, left something behind the building, or went there to collect something. BigFoot is not able to ascertain the identity of the building, but during playback of these tracks, it pauses to alert the operator to a transaction in the temporal vicinity of tracks that are in the spatial vicinity of this building. By viewing this image the operator can ascertain that chocolate milk, a sausage, egg, and cheese "schmuffin," and a Danish pastry were purchased at a Sheetz convenience store (Figure 8). Undoubtedly, many operators would not consider this an interesting signal, but if the operator is an employer, data concerning the subject's whereabouts during working hours may warrant attention, or if the operator is an insurance-company representative, the fact that the subject purchased an unhealthy breakfast may result in increased premiums, mandatory cholesterol checks, and so forth. Similarly, automobile-insurance companies would be very interested in knowing how fast the subject traveled. Because of the high level of their detail, the digital footprint data contain many other examples like these.

Of course, just because an event only happens once does not necessarily make it interesting. Other repeated patterns of activity, such as regularly going to a sports club, bank, or beer store or attending a religious institution, may provide many clues to an individual's persona.

FIG. 9—Tracks recorded between 5:00 A.M. and 12:00 noon on Mondays in State College, Pennsylvania. Differential dot shading reflect different Mondays. Note that, rather than overlaying multiple trajectories from one subject, an operator can overlay the space-time trajectories from multiple digital footprints.

BEYOND THE INDIVIDUAL, A BIGGER FOOT(PRINT)

During exploration of a subject's activities and trajectories, the operator is able to generate and test multiple hypotheses. As each test concludes, more detail—that is, increasingly sophisticated views of behavior as a function of the digital trail—is added to the subject's digital persona. With each new investigation, and indeed each new scenario, analysis techniques become refined to the point where they could be developed into programmable algorithms. For example, if BigFoot were designed to learn from the operator's analysis methodology, it could, theoretically, develop a native form of intelligence that could make the analytical methods extremely powerful, not only for detecting criminal activity but also for many other uses, such as searching for missing persons.

As described here, the system contains only the trajectories and activities of a single subject. However, by adding in additional subjects, commonalities among individuals can be noted, not only in their spatiotemporal behavior but in other activities and interests as well. It is possible to determine in how many spaces two subjects are connected—for example, in space and time, talking on the phone, exchanging e-mail messages, engaging in financial transactions, and sharing journeys. Figure 9 shows several space-time trajectories of one subject moving through space and time on separate days. It was created to look for anomalous behavior; but it also serves as an example of what multifootprint visualization can look like in a system like BigFoot.

Although it may seem obvious that the greater the number of data layers used to represent a digital footprint, the greater the likelihood of uncovering useful associations, it is also similarly likely that entirely coincidental, irrelevant co-occurrences will also be found. Then, because of these coincidences, the likelihood of yet more co-occurrences being found may also increase. The potential for random or erroneous events giving rise to suspicion increases exponentially with every additional subspace or dimension of data in an individual's digital footprint. Furthermore, if someone's digital persona is erroneously tagged as suspicious in some database due to incomplete data, errors in the data, or inappropriate conclusions by an analyst, the onus often falls on the individual to remedy the transgression.

In spaces characterized by fear and suspicion, people's reasonable expectation of privacy—as interpreted by civilians and state authorities alike—diminishes. Moreover, the doctrine of assumed innocence until guilt is proved diminishes, sometimes to the extent that the reverse becomes true. Also, guilt by association in these spaces becomes more prevalent, and the number of coincidences necessary to significantly "color" one's digital persona decreases. Moreover, and consistent with the "law of large numbers," the ability to manipulate digital personae or to misinterpret them increases. Again, we are not calling for the abolishment of existing or developing surveillance technologies. To the contrary, we recognize their necessity; and, furthermore, we support many of their private-sector applications. However, we believe that the concerns raised in this article are heightened spaces characterized by fear and suspicion, so that the tripartite debate becomes increasingly critical.

Any number of scenarios can be invented that raise suspicion resulting from inaccurate time and location data, but with the addition of personal and activities information it becomes much easier to link trajectories to possible explanations—valid or otherwise. The more of this kind of information is gathered, the more we should expect errors of commission (false positives) to arise by chance or through inaccuracies. Hence, without sophisticated models of error and uncertainty, use of this technology is very dangerous. Of course, the converse argument—that a system such as this can help capture criminals or exonerate innocent people suspected of criminal activity—is also substantial. Therefore, legislation concerning the ethical collection and use of digital footprint data is urgently needed. This legislation must be informed by extensive research that seeks to understand a contemporary digital

footprint and to address the accuracy and uncertainty present in the data streams and sources used. For example, researchers must address the following questions:

1. What are the characteristics of all the spaces that collectively constitute a digital footprint?
    a. How are these spaces defined?
    b. How and to what extent do these spaces overlap?
    c. Are some of these spaces static, or do they all constantly evolve?
    d. Who is in control of the shape, content, and evolution of these spaces?
    e. How would these spaces be affected if an individual were to go "off the grid"?
    f. How are these spaces related to commercial or law-enforcement spaces?
2. How is a digital persona different from a true persona?
    a. What sorts of predictions can be made based on digital activities?
    b. What kinds of errors of omission and commission in the system should we consider normal?
    c. In what ways might an informed suspect construct a false digital persona?
3. How can the answers to the above questions inform participants in the debate between privacy and security?
    a. What kinds of intrusiveness are valid in constructing a digital persona, and in what circumstances? (This question might include issues of spatiotemporal scale and issues relating to the activities monitored.)
    b. How do law enforcement and society guard against "guilt by association"?

Clearly, many of the underlying questions are ethical ones. For example, one ethical dimension in surveillance is the good of the many versus the good of the few; put another way, is it reasonable to engage in the surveillance of a few suspects in order to protect many innocent people? A problem with this perspective is that, in fact, we are all subject to the data-gathering activities described above, so what is to stop the few from using such systems to oppress the many?

Appropriate legislation and ethical standards can be defined only after seeking answers to questions concerning these various aspects of a digital footprint. To ignore these empowering technologies and their potential effects on society is potentially very dangerous. Instead, research must be conducted to better understand the technology and its impacts on the wider society.

The sociologist Elia Zureik claims that "Technology is a contingent thing. To put it in the words of Melvin Kranzberg, a leading historian of technology, 'Technology is neither good, nor bad, nor is it neutral' (Kranzberg 1989). Its effects depend on the context of its use" (Zureik 2003, 1). The balancing point, at which the technology that enables good and bad applications of LBS, for example, has not been determined, so the technologies used in BigFoot cannot yet be said to be good or bad. We are concerned about where the balancing point will ultimately rest. Fur-

thermore, as data sets increase in size, so does the likelihood that how one appears—as interpreted from the data—is different from how the subject truly is. Moreover, as data sets grow, multiple, contrasting personas may be generated for one individual at different times, by different people or by different machine algorithms, with varying consequences for the subject. In short, it is not clear—to us, at least—how these developments will ultimately affect how we think of ourselves and how others see us, based on our digital traces.

Discourses abound on all sides of the debates between the desire for new technology, the right to privacy, and the need for security. None of the actors favoring one aspect of this debate is entirely right; and none is entirely wrong. The production of an intelligence infrastructure that simultaneously ensures security and preserves privacy should be the ultimate goal of government. Likewise, with respect to the desire for new technology, provision must be designed into the technologies to ensure privacy where possible, without defeating the potential benefits of their use.

"When there is power, there is resistance, and yet, or consequently, this resistance is never in a position of exteriority in relation to power" (Foucault [1976] 1978, 95). Because resistance and power are entwined concepts, work must be done to democratize technology. A richer understanding of the technology and its impacts on society is the first step toward achieving this goal; and it is to this end that we constructed and tested the BigFoot application: to lay open some of the potential advantages and pitfalls of integrated digital surveillance, to communicate them to other researchers, and to share the application with interested parties.

Notes

1. The same technologies can enable businesses to customize prices based on an individual's profile. For example, an airline could charge higher prices for tickets to customers deemed more likely to pay them and lower fares to customers whose profile makes them less likely to pay more than some threshold amount. A business could thus maximize its profits on a customer-by-customer basis by charging, for each purchase, the maximum amount each customer is likely to pay.

2. A great deal of research is currently under way in this area (see Priyanta, Chakraborty, and Balakrishnan 2000; Snekkenes 2001; Smailagic and Kogan 2002; Barkhuus and Dey 2003; Beresford and Stajano 2004a, 2004b; Hengartner and Steenkiste 2003, 2004; Myles, Friday, and Davies 2003; Barkhuus 2004; Cuellar and others 2004).

3. All of the figures in this article are screen captures from the BigFoot program in action. They are designed for viewing on a computer terminal, not as printed graphics.

4. Researchers interested in experimenting with their own digital footprint are welcome to contact the authors to obtain the latest version of BigFoot.

5. Such information is currently available for a small fee from several online providers. However, we found it to be particularly unreliable because it contained many errors of omission and commission.

References

Baker, S., and K. Baker. 1993. *Market Mapping: How to Use Revolutionary New Software to Find, Analyze, and Keep Customers.* New York: McGraw-Hill.

Barkhuus, L. 2004. Privacy in Location-Based Services: Concern vs. Coolness. Paper presented at the Mobile HCI 2004 workshop on Location System Privacy and Control, Glasgow, Scotland, 13–16 September.

Barkhuus, L., and A. Dey. 2003. Location-Based Services for Mobile Telephony: A Study of Users' Privacy Concerns. In *Proceedings of Interact 2003.* Zurich, Switzerland: ACM Press.

Bench-Capon, T., G. Malcolm, and M. Shave. 2003. Semantics for Interoperability: Relating Ontologies and Schemata. In *Database and Expert Systems Applications, 14th International Conference, DEXA 2003, Prague, Czech Republic, 1–5 September, Proceedings,* edited by V. Marík, W. Retschitzegger, and O. Stepánková. Berlin: Springer-Verlag.

Beresford, A. R., and F. Stajano. 2004a. Location Privacy in Pervasive Computing. *Pervasive Computing, IEEE* 2 (1): 46–55.

———. 2004b. Mix Zones: User Privacy in Location-Aware Services. Paper presented at the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops (PERCOMW '04), Orlando, Fla., 14–17 March.

Ceruzzi, P. [1986] 2000. An Unforeseen Revolution: Computers and Expectations, 1935–1985. In *Technology and the Future,* edited by A. H. Teich, 190–221. New York: Bedford / St. Martin's Press.

Cuellar, J., J. Morris, D. Mulligan, J. Peterson, and J. Polk. 2004. IETF Geopriv Requirements. [http://delivery.acm.org/10.1145/rfc_fulltext/RFC3693/rfc3693.txt?key1=RFC3693&key2=9204649811&coll=GUIDE&dl=GUIDE&CFID=34623567&CFTOKEN=40098573].

Curry, D. J. 1993. *The New Marketing Research Systems: How to Use Strategic Database Information for Better Marketing Decisions.* New York: Wiley.

DHS [Department of Homeland Security]. 2004. *Securing Our Homeland: U.S. Department of Homeland Security Strategic Plan.* Washington, D.C.: U.S. Department of Homeland Security. [www.dhs.gov/xlibrary/assets/DHS_StratPlan_FINAL_spread.pdf].

Dobson, J. E. 1998. Is GIS a Privacy Threat? *GIS World* 11 (7): 34–35.

———. 2000. What Are the Ethical Limits of GIS? *GeoWorld* 13 (5): 24–25.

———. 2002. Geoslavery. Paper presented at the annual meeting of the Association of American Geographers, Los Angeles, Calif., 19–22 March.

Dobson, J. E., and P. F. Fisher. 2003. Geoslavery. *IEEE Technology and Society Magazine* 22 (1): 47–52.

Egenhofer, M. J. 2002. Toward the Semantic Geospatial Web. Paper presented at the Tenth ACM International Symposium on Advances in Geographic Information Systems, 8-9 November, McLean, Va.

Engel, J. F., R. D. Blackwell, and D. T. Kollat. 1978. *Consumer Behavior.* 3rd ed. Hinsdale, Ill.: Dryden Press.

EPIC [Electronic Privacy Information Center]. 2005. Public Opinion on Privacy. [www.epic.org/privacy/survey/default.html].

EPIC and PI [Electronic Privacy Information Center and Privacy International]. 2002. Privacy and Human Rights 2002: An International Survey of Privacy Laws and Developments. [www.privacyinternational.org/survey/phr2002/].

Farmer, D., and C. C. Mann. 2003a. Surveillance Nation: Part One. *Technology Review,* April, 34–43. [http://charlesmann.org/articles/surv-nation.pdf].

———. 2003b. Surveillance Nation: Part Two. *Technology Review,* May, 46–53. [http://charlesmann.org/articles/surv-nation2.pdf].

Fisher, P. F., and J. E. Dobson. 2003. Who Knows Where You Are, and Who Should, in the Era of Mobile Geography? *Geography* 88 (4): 331–337.

Foerstel, H. N. 1991. *Surveillance in the Stacks: The FBI's Library Awareness Program.* Contributions in Political Science, 266. New York: Greenwood Press.

Foucault, M. [1975] 1995. *Discipline and Punish: The Birth of the Prison.* Translated by A. Sheridan. 2nd ed. New York: Vintage Books.

———. [1976] 1978. *The History of Sexuality: An Introduction.* Translated by R. Hurley. New York: Pantheon Books.

———. 1983. The Subject and Power. In *Michel Foucault: Beyond Structuralism and Hermeneutics,* [edited] by H. L. Dreyfus and P. Rabinow, 208–228. Chicago: University of Chicago Press.

Gandy, O. H. 1993. *The Panoptic Sort: A Political Economy of Personal Information.* Boulder, Colo.: Westview Press.

Goss, J. 1995. "We Know Who You Are and We Know Where You Live": The Instrumental Rationality of Geodemographic Systems. *Economic Geography* 71 (2): 171–198.

Guarino, N. 1998. Formal Ontology and Information Systems. In *Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6–8, 1998, Trento, Italy,* edited by N. Guarino, 3–15. Amsterdam: IOS Press.

Hengartner, U., and P. Steenkiste. 2003. Protecting Access to People Location Information. Paper presented at the First International Conference on Security in Pervasive Computing (SPC 2003), Boppard, Germany, 12–14 March. [www.cs.cmu.edu/~uhengart/spc03.pdf].

———. 2004. Implementing Access Control to People Location Information. Paper presented at the 9th ACM Symposium on Access Control Models and Technologies (SACMAT'04), Yorktown Heights, N.Y., 2–4 June. [www.cs.cmu.edu/~aura/docdir/sacmat04.pdf].

Höllerer, T. H., and S. K. Feiner. 2004. Mobile Augmented Reality. In *Telegeoinformatics: Location-Based Computing and Services,* edited by H. A. Karimi and A. Hammand, 221–260. Boca Raton, Fla.: CRC Press.

Hughes, A. M. 1991. *The Complete Database Marketer: Tapping your Customer Base to Maximize Sales and Increase Profits.* Chicago: Probus.

Kashyap, V., and A. Sheth. 1998. Semantic Heterogeneity in Global Information Systems: The Role of Metadata, Context and Ontologies. In *Cooperative Information Systems: Current Trends and Directions,* edited by M. Papazoglou and G. Schlageter, 139–178. London: Academic Press.

Kim, T. J. 2002. Location-Based Services—GIS for Personal Productivity. In *2002 Research Agenda.* University Consortium for Geographic Information Science, Research Priorities. [www.ucgis.org/priorities/research/2002researchPDF/shortterm/b_location_based.pdf].

Klein, M. 2001. Combining and Relating Ontologies: An Analysis of Problems and Solutions. In *Workshop on Ontologies and Information Sharing,* edited by A. Gomez-Perez, M. Gruninger, H. Stuckenschmidt, and M. Uschold, 53–62. Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-2001), 4-5 August, Seattle, Washington.

Kranzberg, M. 1989. The Information Age. In *Computers in the Human Context: Information Technology, Productivity, and People,* edited by Tom Forester, 19–32. Cambridge, Mass.: MIT Press.

Kwan, M.-P. 2000. Analysis of Human Spatial Behavior in a GIS Environment: Recent Developments and Future Prospects. *Journal of Geographical Systems* 2 (1): 85–90.

Ludaescher, B., A. Gupta, and M. E. Martone. 2001. Model-Based Mediation with Domain Maps. In *Seventeenth International Conference on Data Engineering: Proceedings; 2–6 April, Heidelberg, Germany.* Los Alamitos, Calif.: IEEE Computer Society Press.

Lyon, D. 2001. *Surveillance Society: Monitoring Everyday Life.* Buckingham, England: Open University Press.

McCullagh, D. 2004. Database Nation: The Upside of "Zero Privacy." *Reasononline,* June. [http://reason.com/0406/fe.dm.database.shtml].

Michman, R. D. 1991. *Lifestyle Market Segmentation.* New York: Praeger.

Mountain, D., and J. Dykes. 2002. What I Did on My Vacation: Spatio-Temporal Log Analysis with Interactive Graphics and Morphometric Surface Derivatives. Paper presented at GIS Research U.K. (GISRUK): Tenth Annual Conference, University of Sheffield, 3–5 April.

Mountain, D., and J. Raper. 2001. Modelling Human Spatio-Temporal Behaviour: A Challenge for Location-Based Services. Paper presented at Geocomputation 2001, University of Queensland, Australia, 24–26 September.

Myles, G., A. Friday, and N. Davies. 2003. Preserving Privacy in Environments with Location-Based Applications. *IEEE Pervasive Computing* 2 (1): 56–64.

OnStar. 2007. About Us: Fast Facts. [http://209.235.195.221/aboutus_fastfacts.php].

Orwell, G. 1949. *Nineteen Eighty-Four: A Novel.* New York: Harcourt Brace & World.

OTA [Office of Technology Assessment]. 1987. *The Electronic Supervisor: New Technology, New Tensions.* Washington, D.C.: Congress of the United States, Office of Technology Assessment.

Pickles, J. 1995. Representations in an Electronic Age: Geography, GIS, and Democracy. In *Ground Truth: The Social Implications of Geographic Information Systems,* edited by J. Pickles, 1–30. New York: Guilford Press.

Priyanta, N. B., A. Chakraborty, and H. Balakrishnan. 2000. The Cricket-Location Support System. Paper presented at the Sixth International Conference on Mobile Computing and Networking, Boston, 6–11 August.

Raper, J., D. W. Rhind, and J. W. Shepherd. 1992. *Postcodes: The New Geography.* Essex, England: Longman Scientific and Technical.

Robbin, J. E. 1980. Geodemographics: The New Magic. *Campaigns and Elections* 1 (1): 25–34.

Rule, J. B., D. McAdam, L. Stearns, and D. Uglow. 1980. *The Politics of Privacy.* New York: New American Library.

Sheth, A. P., and J. A. Larson. 1990. Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases. *ACM Computing Surveys* 22 (3): 183–236.

Smailagic, A., and D. Kogan. 2002. Location Sensing and Privacy in a Context-Aware Computing Environment. *IEEE Wireless Communication* 9 (5): 10–17.

Snekkenes, E. 2001. Concepts for Personal Location Privacy Policies. Paper presented at the Third ACM Conference on Electronic Commerce, Tampa, Fla., 14–17 October.

Taylor, H. 2003. Most People Are "Privacy Pragmatists" Who, While Concerned about Privacy, Will Sometimes Trade It Off for Other Benefits. *Harris Poll,®* No. 17, 19 March. [www.harrisinteractive.com/harris_poll/index.asp?PID=365].

Thomas, R. K., and R. J. Kirchner. 1991. *Desktop Marketing: Lessons from America's Best.* Ithaca, N.Y.: American Demographic Books.

Weaver, S. D. 2005. The Visualization of My Digital Footprint: A System to Evaluate the Societal Consequences of Ubiquitous Computing with Emphasis on Homeland Security and Privacy Erosion. M.S. thesis, The Pennsylvania State University. [www.geovista.psu.edu/publications/2005/weaver_MSthesis.pdf].

Weiser, M. 1991. The Computer for the Twenty-First Century. *Scientific American,* September, 94–104.

Zureik, E. 2003. Who Knows What about Whom? Paper presented at the Third UNESCO Philosophy Forum, Who Knows What? Paris, 14 September.