Database Risk Management in Digital Mapmaking Company---Navteq

Rob Chapman *

Ying Sai *

Abstract

The focus of this study is on map database risk management. Map databases are the foundation of automotive navigation systems. It is a part of the satellite navigation system designed for use in automobiles. While map database management in many ways is similar to general business database management, it also has its unique set of challenges. Securing the map databases while updating and changing is one of the most critical challenges in digital mapmaking. This research demonstrates how a world leader in the industry had adapted a risk management approach to analyze potential risks and took the necessary security measures to reduce those risks. This research contributed to the information system field by introducing a pragmatic approach in risk management and demonstrated its effectiveness in database management problem solving.

Introduction

This study examines the information security, especially the database management and security of a digital mapmaking company. NAVTEQ is a Chicago based provider of Geographic Information Systems (GIS) data and is a dominant company in providing the base electronic navigation maps. The company is a wholly owned subsidiary of Nokia but operates independently (NAVTEQ).

NAVTEQ's underlying map database is based on first hand observation of geographic features rather than relying on official government maps. Its maps currently cover 72 countries on 6

Rob Chapman, *College of Business Administration,Loyola Marymount University, 1 LMU Drive, Los Angeles 90045, California USA*

Dr. Ying Sai, *College of Business Administration, Loyola Marymount University, 1 LMU Drive, Los Angeles 90045, California USA, ysai@lmu.edu*

continents, and new coverage is being added each day. It provides data used in a wide range of applications, including automotive navigation systems for BMW, Chrysler, Mini and many other carmakers (accounting for around 85% of the market sales share). It also produces customized map databases for in-vehicle navigation systems used by drivers or in mobile phones throughout North America and Europe.

Portable GPS devices made by Garmin, Magellan, Lowrance and web-based applications, such as Google Maps, Yahoo! Maps. Local Live and MapQuest also use its maps. Microsoft's Flight Simulator X uses NAVTEQ data for automatic terrain generation.

NAVTEQ maps offer accurate road geometry along with up to 260 road attributes—from turn restrictions to road construction barriers and speed restrictions. Plus, the NAVTEQ digital map database contains millions of Points of Interest: everything from restaurants and stadiums to hotels, fuel stations, and hospitals.

The tremendous geographic scope of the company's products combined with the complexity of the maps themselves combine to create a more opportunities for security risks to cause problems. NAVTEQ has attempted to minimize or contain several of these risks.

Navteq Map Database

The main asset that the company is trying to protect is the product that the company licenses to customers--the map database itself. There are other physical assets at risk, such as computer hardware, GPS receivers, video collection equipment, and other hardware. NAVTEQ also maintains a fleet of vehicles used for data collection, most of which also contain expensive data storage equipment. However, these items are covered by insurance and the loss of any single piece of hardware would not be catastrophic. On the other hand, the map database is what brings in almost all revenue to the company, so the focus of this study will be on the risks to map database.

The map data is extracted into many different formats such as formats made specifically for in-vehicle navigation. These formats include SIF (Standard Interchange Format) and GDF (Geographic Data File). Other versions are produced for GIS software such as ESRI's Arc products and for hand-held devices that enable pedestrian routing by including sidewalks, etc.

Each of the databases will cover only one country or a portion of a country. For example, the NAVTEQ map of the United States is made up of ten separate databases each containing several states, which is done in order to prevent file size from becoming unmanageable for NAVTEQ and its customers.

NAVTEQ is also trying to protect several archived versions of every database. The standard release cycle for NAVTEQ product is quarterly, plus there is always a current unreleased version that is still being worked on. A customer may decide to license the third quarter 2007 map, but may spend nine months testing the product and compiling it for use in their systems. At any time, if a large enough problem arises, NAVTEQ may be required to retrieve an archived copy of the database, make corrections, and then re-release it to the customer. The simple math of about 100 coverage areas, multiplied by up to 49 different types of data extracts per area, multiplied by at least 4 previous quarters of released data equals a very large number of databases to protect.

NAVTEQ has other assets to protect in addition to its core product. One of these assets is NAVTEQ proprietary GIS software called Atlas. This software was developed in-house and has been continually refined since the company was founded in 1985. It is what employees use to convert geographic information from field data collection, government sources, phone verified data, and other areas into the map product that NAVTEQ sell to its customers. If the Atlas software were somehow tampered with, it could lead to data errors. The software and all subsequent updates are distributed to all users automatically, through a download program called Marimba. The software only goes to machines that are recognized by the download software. The Atlas software can only be used by these same machines and then only while connected to the company network, either in an office or remotely through VPN.

NAVTEQ also values its database creation methods and work processes such as the use of GPS to collect positional data along with aerial imagery and other sources, which the company believes gives them a competitive advantage (Fig1). However, the scope of this paper will limited to the security of the map database.

Fig 1. Automobile navigation device with GPS

Finally, NAVTEQ is trying to protect the image and reputation of the company. The image that its customer perceives is important, because they are ultimately the ones who choose to license NAVTEQ data for their navigation products. They are responsive to data errors that are found through their own testing or by end users. Customers also judge NAVTEQ on how well the company manages risk, and the processes that have been put in place to reduce that risk. NAVTEQ also wants to maintain a positive reputation among end users of navigation products that use its data. Most customers are believed to select their navigation system based on the features and price of the system itself, not the map database installed on it. However, as these systems are better known, the names of the data providers could become a larger factor in the buying decision.

Challenges in Map Database Management

Databases are the most valuable and most critical resources NAVTEQ has. Its databases are used by thousands of customers and databases are critical part of the customer's operation. In general, the more encompassing the database, the more systems and business functions it touches, the greater utility of the database. At the same time, the more encompassing the database is, the greater the potential for problems.

The responsibility of a database administration (DBA) can be categorized into four sections: development, operation, backup/recovery and adaptation. In this study, authors mainly analyze the security risk after the development stage. (Knoenke 2007)

Map database management has its unique characteristics. As automotive navigation systems are playing an increasingly important role in the emerging areas of Location-based services and advanced driver assistance systems, the real-time related requirement for an onboard map database become ever more important. Maintaining such a map database, including keeping it up to data and incorporating accurate related information is the most challenging part of NANTEQ's business.

Content of a map database

A map database represents a road network along with associated features. A road network comprises basic elements (notes, links and areas) and properties of such elements (location coordination, shape, addresses, road class, speed range, etc). Other information associated with the road network is also included, such as points of interest, building shapes and political boundaries (Fig 2).

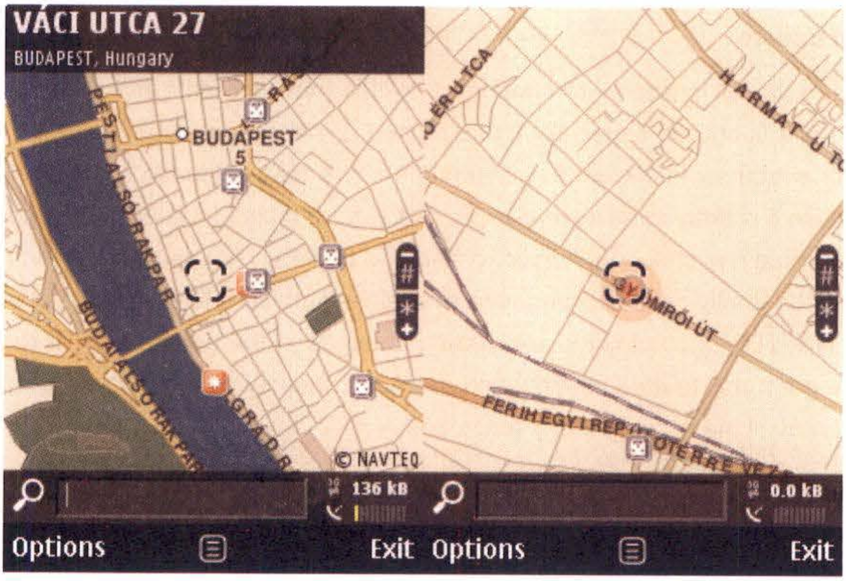

Fig 2. An example of NAVTEQ map

Interchange format

NAVTEQ collects, aggregates and supplies data in a well-defined file format that is specifically intended for information interchange, i.e. Standard Interchange Format (SIF). A small number of record types are used to represent the various types of data and each record type consists of sequence of fields, which are either fixed length or delimited.

To turn such database into a runtime format that used in auto navigation systems, mapmaker will have to work with the automaker to go through a compilation process, which includes the

following five steps (Longley 2005):

1. Check for network consistency. For example, ensure that all no tes that should be connected by a link do have such link.
2. Assign identifiers (IDs) to all entities in a systematic manner.
3. Apply multiple sets of indices to entities to facilitate searching the database in an expected ways.
4. Replace multiple occurrences of data items by indices into tables containing a single copy of each such item.
5. Apply compression techniques to reduce the overall size of the database.

Incremental update

For most navigational functions it is important to have an upto date map database, and for some functions it is critical, especially those related to safety and public security. It would be impractical or extremely inefficient to transmit the entire new database to replace an existing version because it is likely to be several gigabytes in size. Instead it is desirable to transfer just the information that has changed. A major difficulty is that any change made to the content of a map database generally causes changes to all assigned entities. To maintain a map database, over half of NAVTEQ 's employees are making changes to the database on a regular basis. This poses a great risk to data integrity and database security. In the following section, an in-depth analysis of all the security risk is presented.

Analysis of Risks in Map Database Management

The risks to the databases are internal and external, accidental and intentional. The most common way that the database is damaged is through unintentional errors introduced to the database by employees. NAVTEQ has over 1,600 employees, of which over half are making changes to the database on a regular basis. More than a risk, it is an absolute certainty that errors will be introduced. The more specific risk is of an error being created that a customer would view as being a degradation of NAVTEQ product; an error that they would want corrected and incorporated into a reshipped product. The problem is that it is not always easy to predict which errors a customer considers critical. Some customers' compilation programs, which are used to convert NAVTEQ data to a format specific to their hardware, can crash upon encountering a seemingly innocuous problem.

The chance of an employee intentionally damaging the database is much less likely, but still cannot be discounted. A person could delete large sections of the map from an area, or be subtle and make attribute changes that are technically correct, but do not reflect the real world that the map represents. A disgruntled employee may try to do something to embarrass the company, such as inserting profanity into the database. Many intentional coding errors would be caught before the product was released by the same software used to find accidental coding mistakes. The rest can be partially mitigated by carefully controlling access to the database itself.

There is also the risk of large-scale damage to the database itself, such as data being corrupted or overwritten. Companies that license or sell data have a unique vulnerability because there really is only one true copy of the entire product and a series of backups. It is a lot more likely that something would happen to a database than it would be for a car company to lose its entire inventory of cars for example. Including back-ups the databases are stored on 96 servers, with most running the Sun Solaris operating system. All of these servers are dedicated for map database use only, thus avoiding the risk of them being corrupted by other functions.

Besides human errors, there could be errors generated by sensor devices. One of the common errors is generated by imperfect GPS signals that caused by buildings (Chen, Li and etc 2005), or other technical problems like backward progress and local minimum (Kwon and Shroff 2006).

The risk to the reputation and image of the company by its customers and the end users of the customer's products must also be considered. The corporate customers, such as handheld GPS manufacturers and automotive electronics firms have their own testing benchmarks to judge the quality of the map data. All of the measures that NAVTEQ puts in place to reduce risks to the database also serves to improve the results of customer testing. On the other hand, the perception that end users have of the map database is based on their own experiences with navigation products, but also on sensationalized media accounts of database inaccuracies.

Two recent examples of map error related accidents that occurred while drivers were using navigation systems are highlighted below. The implication in these types of stories is that an error in the map database was responsible for the accident, and they usually convey the general attitude that this type of thing did not happen when people relied on paper maps.

"The driver of the bus carrying the Garfield High School girl's softball team that hit a brick and concrete footbridge was using a GPS navigation system that routed the tall bus under the 9-foot bridge, the charter company's president said Thursday. The off-the-shelf navigation unit had settings for car, motorcycle, bus or truck. Although the unit was set for a bus, it chose a route through the Washington Park Arboretum that did not provide enough clearance for the nearly 12-foot-high vehicle. The driver told police he did not see the flashing lights or yellow sign posting the bridge height." (Koutsky 2008)

"A Global Positioning System can tell a driver a lot of things — but apparently not when a train is coming. A computer consultant driving a rental car drove onto train tracks Wednesday using the instructions his GPS unit gave him. A train was barreling toward him, but he escaped in time and no one was injured. The driver had turned right, as the system advised, and the car somehow got stuck on the tracks at the crossing. He jumped out and tried to warn the engineer by waving. He got out of the way just before the train slammed into the car at 60 mph, Metro-North railroad spokesman Dan Brucker said Thursday. The car was pushed more than 100 feet during the fiery crash. Some 500 train passengers were stranded for more than two hours during the Wednesday evening rush hour. The accident also heavily damaged 250 feet of rail, Brucker said."(MSNBC 2008)

Both of these cases, while amusing to readers, serve to harm the image that the public has of NAVTEQ and its products. This in turn can also cause an overreaction by corporate customers, which might want an investigation of the issue, even if the map itself was not at fault. The first place that reporters turn to when writing one of these news reports about a "navigation system accident" is to the company that sold the navigation system. It is easy for the hardware manufacturers to turn around and blame the map provider instead of addressing the real problem, which is often driver error.

This risk is difficult to control, because NAVTEQ has little control over the articles that are written that reference its map products. However, the company can take a few steps to reduce the damage. First, the company can put forward a spokesperson to voice the company's side of the story. These people stress that any navigation system is meant to be used as a guide, and should not be followed blindly. In the case of the bus hitting the bridge, the driver of the bus was responsible for knowing the height of his bus and comparing it to the warning signs posted on the bridge. This is common sense to most people, but having a spokesperson reiterate it helps people understand that the navigation system did not "make" the bus hit the bridge.

There are other ways that the company can minimize the reputation damage from these high-profile incidents. The most obvious one is to minimize any map errors in the first place. One way this is done is to give both customers and end users the ability to report map errors to the company. These are then acted on in timely manner, which can help prevent these incidents in the first place.

Solutions to Mitigate Risks

The risks caused by employee errors are mitigated through training of proper database specifications, geocoding procedures, and through the use of quality checks. However, the database is so complex that training would be totally insufficient in preventing errors from being introduced by employees. The main tool NAVTEQ uses to mitigate the risk is the database software itself. Currently, the company's map interface, Atlas, uses 1,053 different validations to flag potential coding mistakes. These validations create a database exception when a potential error is created, for errors ranging from section of road geometry that does not connect to the network to a direction of travel that has been reversed. Users can then fix the error and mark it as legitimate, or leave the exception to be fixed later.

While the danger of mistakes in the database caused by employee error may not carry the excitement of an attack by hackers, the risk to the reputation and bottom line of the company is just as great. One error, if it is severe enough, could cause customers to demand a reship of the product. If the problem is caught by the customer, and it is for an in-vehicle navigation system, there may not be time for them to re-compile the product in time to get the DVDs to the assembly line. Imagine how upset a high-end car manufacturer would be after marketing the navigation system as an expensive (and high margin) option--only to have the map data contain errors or be out of date. Overall, the Atlas software is very effective at mitigating the risk of user generated data problems, which has lead to greater customer satisfaction as the system has been refined over the years. The system is not perfect, however, and as new content is added (such as stop lights and evacuation routes), new validations are needed to catch the errors related to these attributes.

User Maintenance Tool

One part of the security solution NAVTEQ developed to reduce the impact of accidental errors or employee sabotage of the database is to restrict access. This is done via an interface the company calls the User Maintenance Tool. This system allows each employee to be granted permission to integrate changes to the database by a specific geographic area, down to the county level if needed. A new employee is given access only to the counties where he or she is working, thus containing any intentional database damage, but not preventing it. This method is more effective than it sounds, however, because areas where work is ongoing would be subject to many levels of review, increasing the likelihood that the problem would be found. A cunning saboteur with access to the entire database would go to an area where no work was taking place for the quarter, thus increasing the risk that the problem would not be detected by quality checks.

There are, however, several security loopholes that the User Maintenance Tool leaves open. First, changes to a user's access level are not tracked in any way, and the access levels are not reviewed for accuracy. This means that there is no way to spot trends, such as new employee who suddenly had the access of a system administrator. Secondly, any user can grant access to any user below them. Ideally, these types of changes should be limited to the employee's supervisor, and maybe a handful of others. The procedure for either increasing an employee's security level or regions they have access to usually involves a phone call or email from the user stating that they need increased database privileges. The reasons why are rarely questioned, and an atmosphere of trust exists which could be exploited by an unscrupulous person.

Change Tracking

Another tool that the company has to contain problems, both accidental and intentional, is change tracking (Ludtke 2008). Every change made to the map database is logged, stored, and can be queried upon later. If it is found that an employee or group of employees are creating errors, it would be possible to trace each change made to that user's ID number and undo those changes if needed. Once the problem is traced back to a particular user or group of users, then any commonalities in the group can be investigated. For example, were all participants in a particular training class miscoding a particular attribute of the database? If so then clarifying instructions could be sent out get at the root of the problem. This system works well for containing coding mistakes if the Database Engineer chooses to investigate the change and tries to find out the root cause. One way that the change tracking system is lacking is that it is not currently possible to revert or undo a change once it has been integrated into the database. The tool says specifically what change was made, when, and by whom, but the user or someone else still has to go in and manually correct the problem.

There was a case in February 2007, which demonstrated the usefulness of this tool. It was discovered during a statistical comparison between the Q4-07 product and the Q1-08 product that some highway signs had been removed from the database. First a query on all signs was run in the areas that showed a significant decline. The company has set percentages of increase or decrease within a quarter that are considered suspicious and should be identified. From the query output, signs that existed in the Q4-07 database but not in the Q1-08 database were investigated using the Change Tracking Tool. The tool showed which employee made the error and when, and soon it became apparent that the problem rested with one group of employees. Further investigation revealed that crucial instructions had been misunderstood, which led to a clarification in their training and a repair to the database before it was released. The combination of database access restriction, change tracking and database validations cannot remove 100% of the risk, but when used together they come very close. As long as people are making changes to the database there is always a chance that error will be introduced.

The risk of an entire database being deleted, corrupted, or in any way made unusable is alleviated by a system of using replication servers to continually create backup copies of all databases. For example, November 2006 the entire Q4-2007 North American database and all accompanying log were overwritten because a program designed to create a new database was accidentally run. Because all active databases are backed up internally throughout the day, the map was restored to an earlier version and only about 1 hour's worth of work was lost. When this loss is multiplied across hundreds of employees, the cost was significant, but much less so than if work was lost for a whole day, week, etc.

In many scenarios an internal backup would not be sufficient because of a natural disaster or other catastrophe could destroy all on-site back-up systems. For those instances what is referred to as a hot-replicated system is in place; data is stored at the Fargo, ND production facility as well as the headquarters in Chicago. This type of system allows the back-up data to be used almost immediately. As a final redundancy, copies of all databases are sent to an off-site data storage provider, Iron Mountain. These copies are made daily and sent to Iron Mountain so that no more than one day's work could ever be lost.

The software used to update the map database, Atlas, currently has 1,053 validations or rules, which run against the database and all changes to it. Some rules are run live and others run only as a batch program, depending if database context is needed to properly evaluate the situation. A unique risk with this system is that users have a tendency to trust the software too much. If the system indicates an error many users will try correct the problem (even if none exists), in effect creating an error where none existed before. This can be mitigated somewhat by thoroughly testing new rules and providing proper documentation to users, but this does not catch the problem in every case.

Trade-offs in risk management

From a user's perspective, the software that limits access to different databases, the User Maintenance Tool, can greatly slow down the workflow because of the time it takes to gain access to a new database. Finding a person to grant access, waiting for them to make the change and integrate it, and all of the communication back and forth takes up valuable time that could be spent improving the product. Some work may be lost if the user has made changes to the database without realizing that he or she did not have the proper access need to integrate the changes.

Another trade-off is in the interface to the map database, Atlas. As more and more validations are added to flag the untold combination of possible errors, the system performance can suffer. An update to the software that closes a loophole somewhere may cause the software to lock up or crash. Even though all software changes go through User Acceptance Testing, it is difficult to simulate the load of all users on the system at the same time.

Conclusion

This study focused on map database risk management. A map database is the foundation of automotive navigation systems. It is a part of the satellite navigation system designed for use in automobiles. It typically uses a GPS navigation device to acquire position data to locate the user on a road in the map database. Using the road database, the device can give directions and to provide location based services such as finding the nearest bank or gas station.

While map database management in many ways is similar to business database management, it also has its unique set of challenges. In this paper, the authors have identified these challenges as special content of map database, interchange format and incremental update. Due to its unique characteristics, map database management also involves a special array of risks that can be categorized into two dimensions, such as internal/external and accidental/intentional. This paper has provided a detailed analysis for the risks and also suggested well-documented and proven solutions that will mitigate those risks.

Securing map databases while updating and changing take place has become a major challenge in digital mapmaking industry. This research demonstrates how a world leader in the industry had adapted a risk management approach to analyze potential risks and took necessary security measures to reduce those risks. This research contributes to the literature in information system security by introducing a pragmatic approach in risk management with demonstrated effectiveness in map database management.

References

1. Chen, W, Li, Z. Yu, M and Chen Y. Effects of sensor error on the performance of map matching, *Journal of Navigation*, 58:2: 272-282, 2005
2. Holzwart, J. Incident Management Plan, NAVTEQ Internal Document, September 2007

3. Kavuri, P. Map Infrastructure Database Copy Process, NAVTEQ Internal Document February 2008
4. Koutsky, D. Garfield High School softball bus crashes loses roof, http://seattlest.com April 17th, 2008
5. Kown, S. and Shroff, N.B. Geographic routing in the presence of location errors, *Journal of Computer Networks*, 50:15:2902, 2006
6. Kroenke, D. *Using MIS,* 2nd edition, Pearson Prentice Hall, 2007
7. Lacko, S. Oracle 10g Upgrade Project Scope, NAVTEQ Internal Document, February 2008
8. Longley, P. *Geographic Information Systems and Science,* Books.google.com, 2005
9. Ludtke, D. *Update of file-system-based navigation databases*, paper presented at IEEE Vehicular Technology Conference, 2008
10. MSNBC, Man using GPS drives into path of train, MSNBC technology and science, tech and gadgets, Jan 3, 2008
11. NAVTEQ www.navteq.com
12. Sullivan, P. IT Disaster Recovery Plan, NAVTEQ Internal Document, May 2006
13. Vilensky, I. Inventory of Oracle Hosts and DBA Assignments, NAVTEQ Internal Document, January 2008