School of Computing
FACULTY OF ENGINEERING

UNIVERSITY OF LEEDS

Buried Utility Pipeline Mapping Based on Multiple Spatial Data Sources: A Bayesian Data Fusion Approach

Huanhuan Chen and Anthony G Cohn

School of Computing
University of Leeds

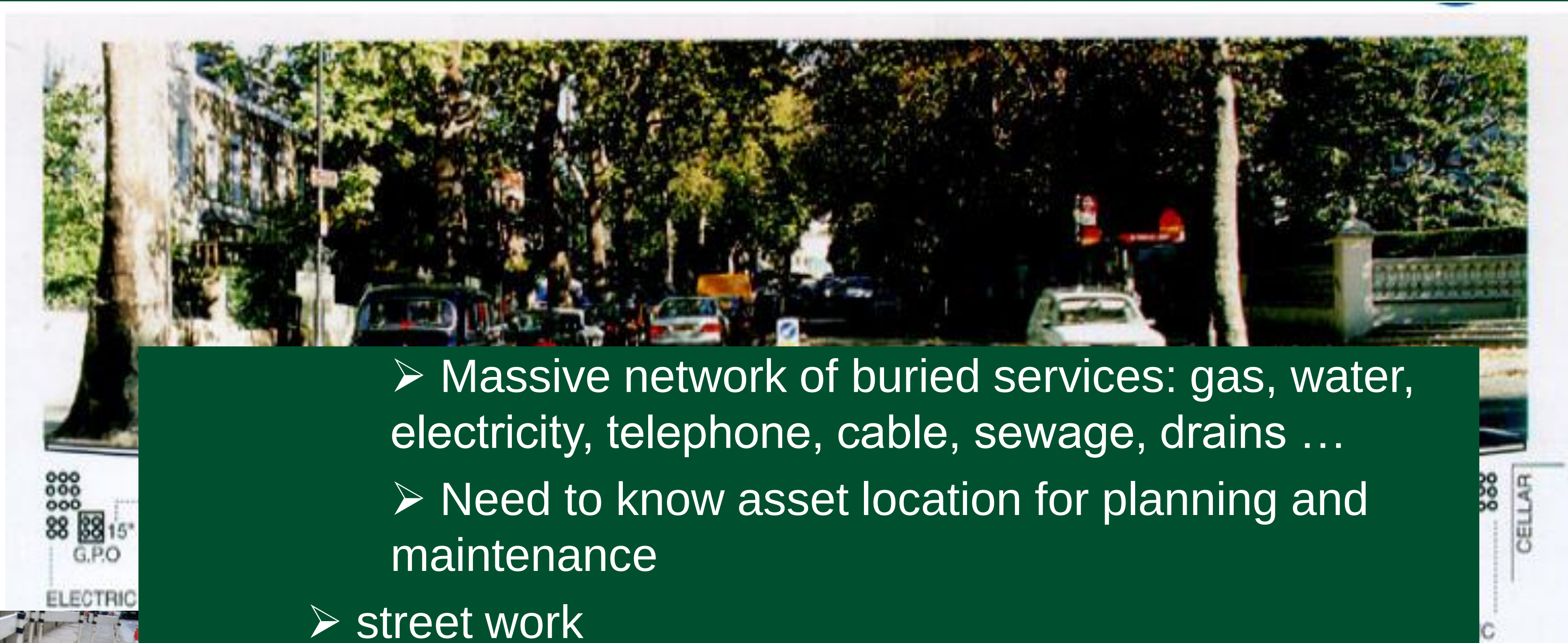
Background

Utility infrastructure has been part of the urban fabric for millenia

Modern utility services established in the 19th Century

- *Most assets laid in the street*
  - *Which has become increasingly congested*
- *Organisation\legislative structures have changed dramatically*
  - *Private – public – private companies*
- *Asset recording systems have developed autonomously*
  - *Until mid 80s based on hand drawn engineering plans*
  - *Since 80s many companies have migrated to GIS based asset records*

The Problem



- Massive network of buried services: gas, water, electricity, telephone, cable, sewage, drains …
- Need to know asset location for planning and maintenance

- street work
  - ~4M street openings p.a.
  - Direct costs of £1B p.a.
  - Indirect costs of £3B-£5B p.a.
  - Safety!

Don’t the utility companies have maps?

UNIVERSITY OF LEEDS

Incomplete

- lacking depth information
- missing records
- missing attributes

Inaccurate

- spatially relative to features no longer present/visible
- errors introduced through digitisation
- inaccurate survey technologies
- deliberate inaccuracies (for better display)

Can’t we use sensors to find the buried assets?

UNIVERSITY OF LEEDS

- Expensive in time
- Ground Penetrating Radar (GPR) doesn’t work in clay
- Plastic hard to detect

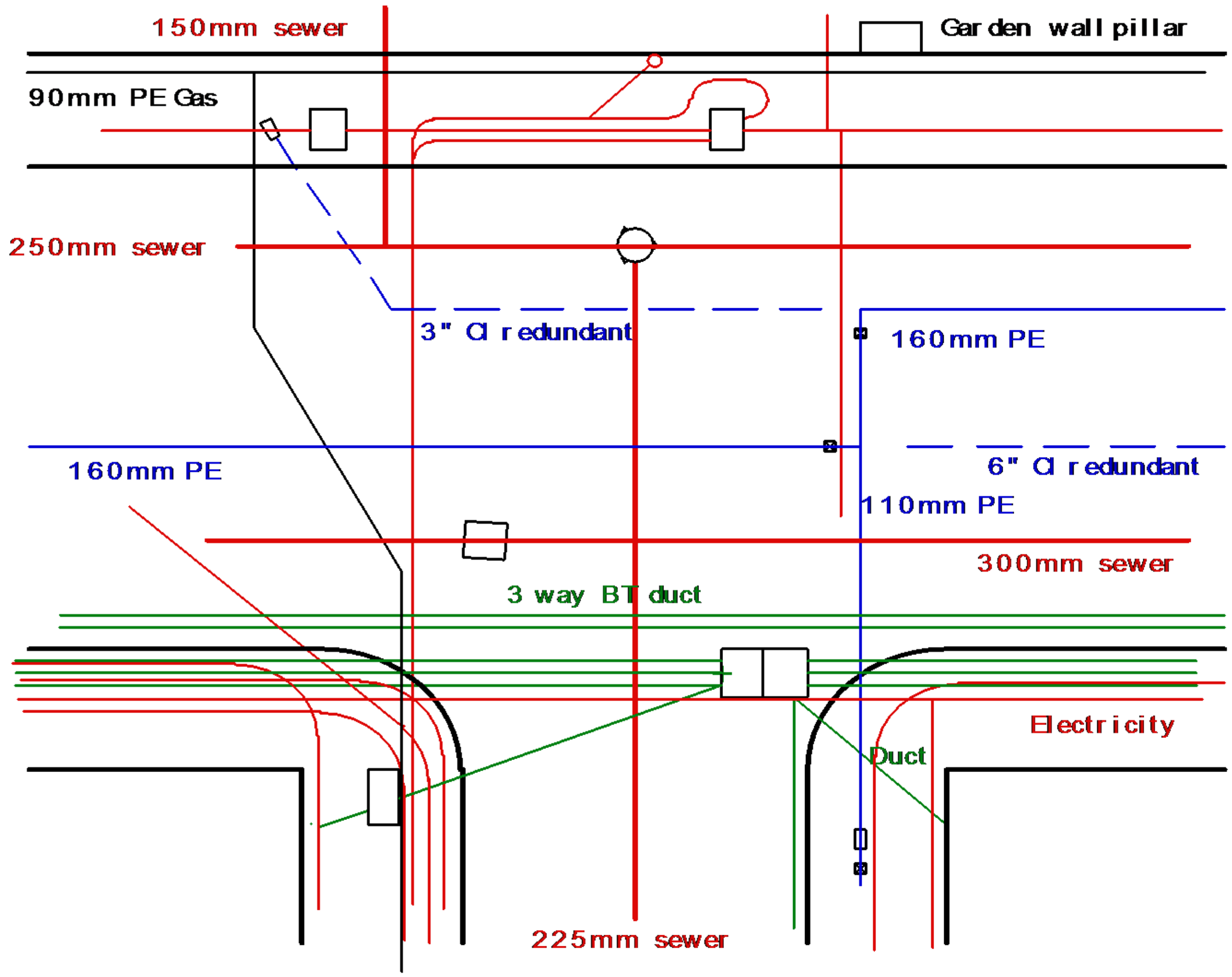
Ground truth

UNIVERSITY OF LEEDS

BURIED ASSETS KNOWN TO BE BENEATH THE SURFACE …

… As uncovered during a complete reconstruction of the junction

(UKWIR Location Trials in 2001)



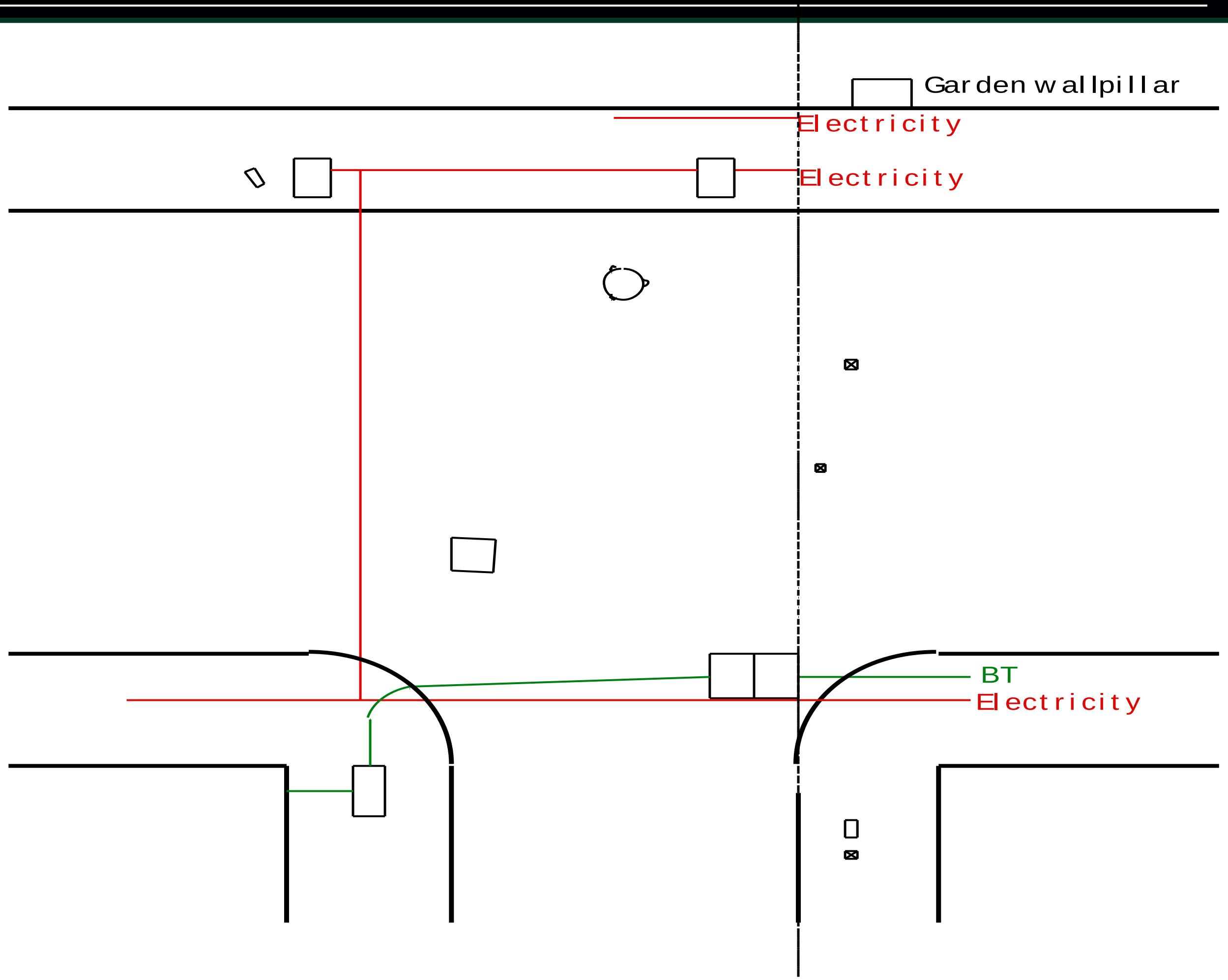
One surveyor's map:

UNIVERSITY OF LEEDS

PLAN OF BURIED ASSETS …

… As produced by one of three survey contractors

(UKWIR Location Trials in 2001)



A new approach

- Multi sensor
  - Ground Penetrating Radar (GPR)
  - Sonar sensor
  - 2 x Magnetic field sensors
- Use expectations from utility records
- Intelligent data fusion

The data fusion problem

UNIVERSITY OF LEEDS

How to build a map from sensor data?

Two approaches:

1) Analogical approach: treat a map as pixel array (raster image)

- no explicit knowledge of what

2) Symbolic approach: vector map

– each asset is vectorised polyline

- can have symbolic attributes (e.g. size, material...)

Modern Utility records are in second form

Approach at Leeds is to create symbolic, vectorised map.

Need to convert all sensor readings to a symbolic hypotheses

- e.g. There is an asset, at position x,y,z, of diameter d

  - hypotheses can have probabilities.

Inputs to data fusion

UNIVERSITY OF LEEDS

1) Statutory Records

- VISTA: Visualising Integrated Information on Buried Assets to Reduce Streetworks (£2.4 M)

- Integrate and Visualize Heterogeneous Utility Data from different Utility companies (24 Partners)

- Syntactic and semantically homogenous vectorised, attributed map

2) Survey of Street furniture

- x,y,z locations of each manhole

- estimate directions of assets as they leave manhole

- depth of assets in manhole

3) Sensor data

- GPR, acoustic, EM...

Currently just GPR

The problem

Statutory records are

- incomplete (may not record all apparatus, or not all attributes, or only x,y, not z)
- inaccurate (may record items in wrong absolute position)
- inaccurate (may have wrong attribute information)

Sensor readings may be:

- noisy (sense things that are not apparatus)
- incomplete (may not detect everything (soil, asset material))

How to “join the dots” – i.e. link up readings from different points

Approach: Bayesian Data Fusion (BDF)

Treat problem as *finding the most probable interpretation,* given the evidence:

1. the statutory records
2. survey of street furniture

   (and approximate direction of pipes after lifting cover)
3. the sensor readings
4. information about soil

   - ...

Statutory Records provide a Prior Probability on asset locations

UNIVERSITY OF LEEDS

Statutory records are inaccurate and incomplete.

For each sensor detection, can search utility records for possible matches, to give a prior probability that the detection relates to that record.

Street furniture Data

a series of manhole locations and estimated asset directions, along with an uncertainty matrix.

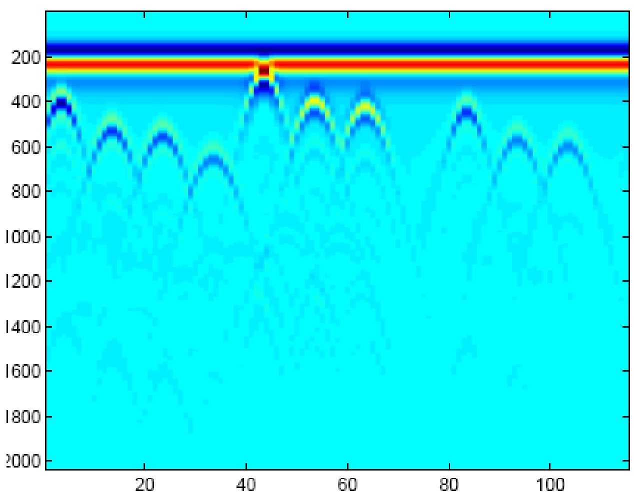
Hyperbola Identification

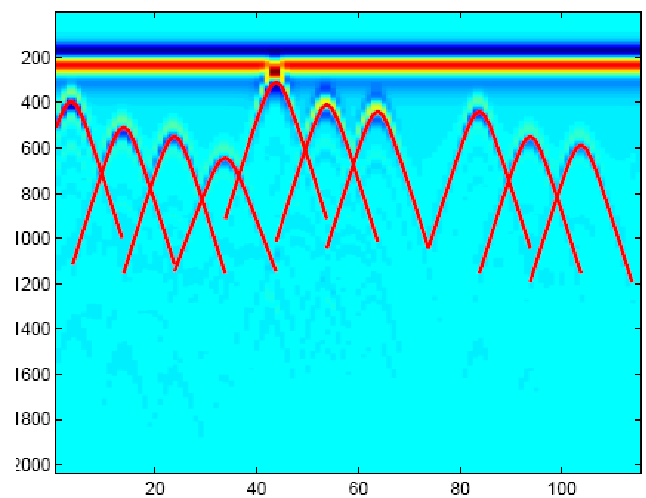
UNIVERSITY OF LEEDS

Automated algorithm to find hyperbolae in GPR scans

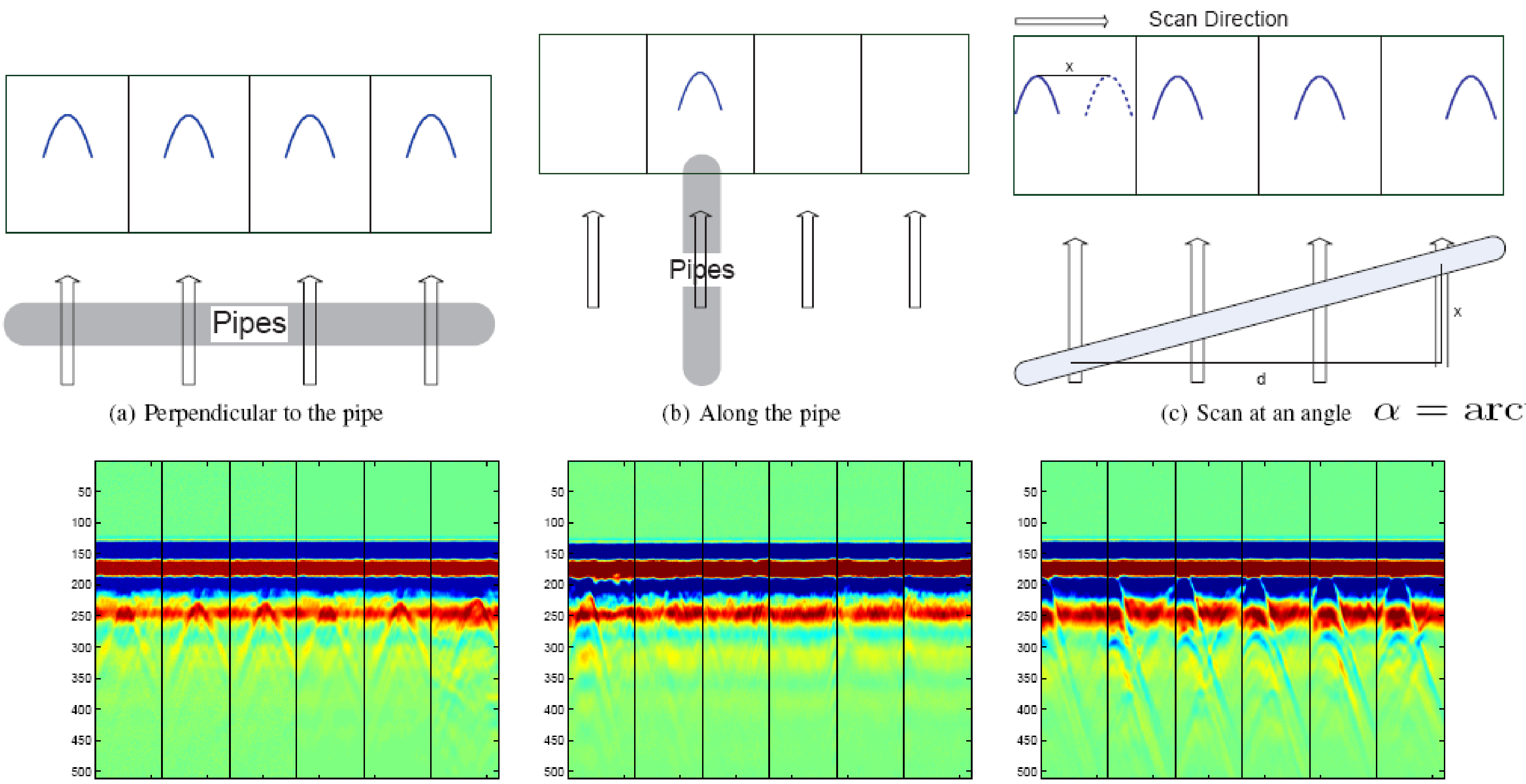
Problems:

- Noise in data
- How many hyperbolae?
- Real time operation desirable





Find hyperbolae in successive parallel GPR scans: estimate direction



(a) Perpendicular to the pipe

(b) Along the pipe

(c) Scan at an angle $\alpha = \arctan \frac{x}{d}$

(d) GPR data: perpendicular to the pipe

(e) GPR data: Along the Pipe

(f) GPR data: Scan at an angle

Sensor data (GPR)

- Find hyperbolae in scans

- [Chen and Cohn, 2010a,] outlines the approach to determine the approximate direction of a pipe segment.
- Each observed pipe can be represented by a vector ***O*** = (x, y, Ɵ) and the uncertainty is represented by a matrix **Co**.

O = (x, y, Ɵ)

$\mathbf{C_o}$ = diag(Δx,Δy,ΔƟ)

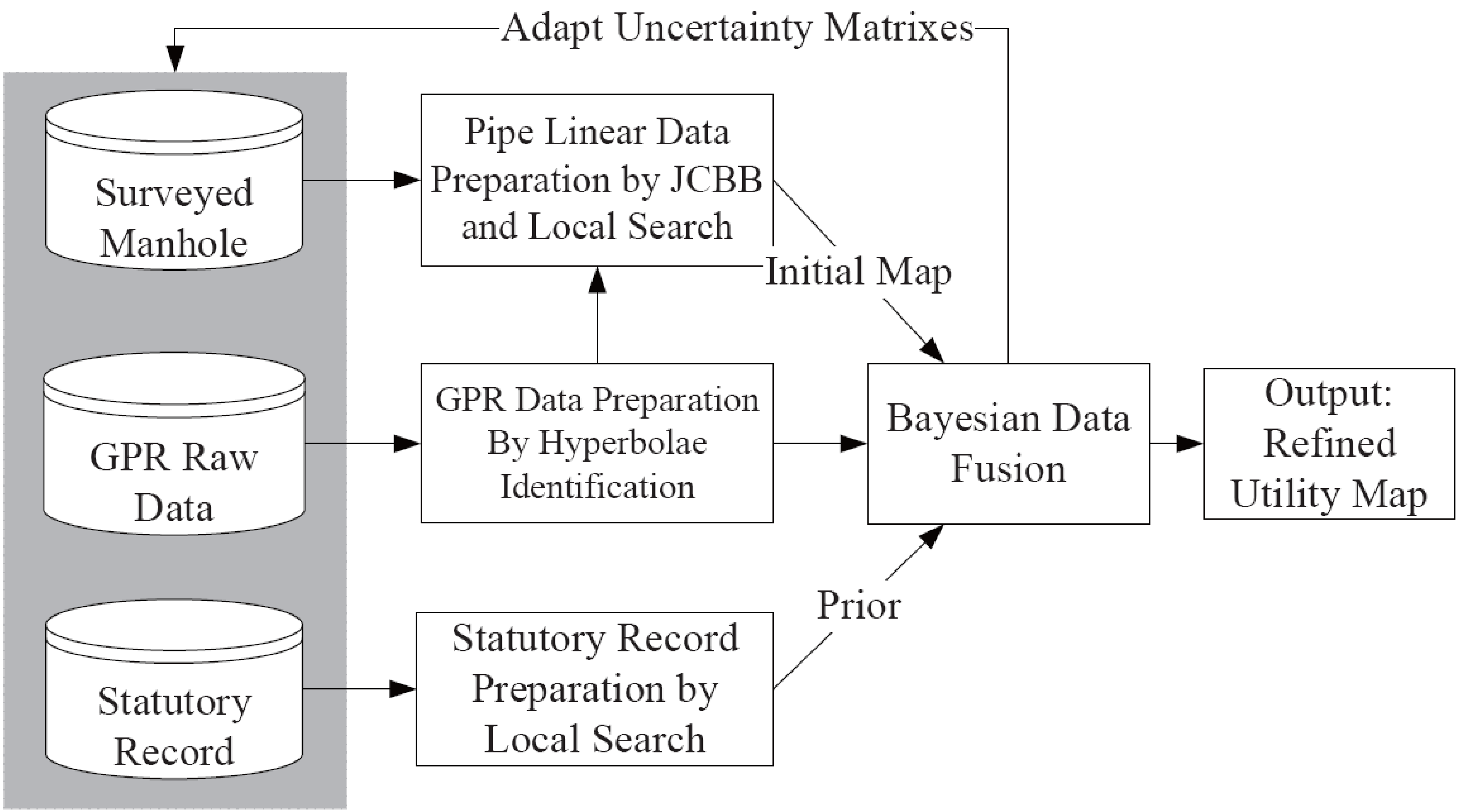
Pipe linearity assumption *L*

Pipe linearity assumption is an indirect data source based on the observations of other pipe locations.

- Most pipes are approximately linear, this variable provides evidence as to whether to connect pipes segments detected by GPR scans or manhole inspections.

Flowchart of BDF

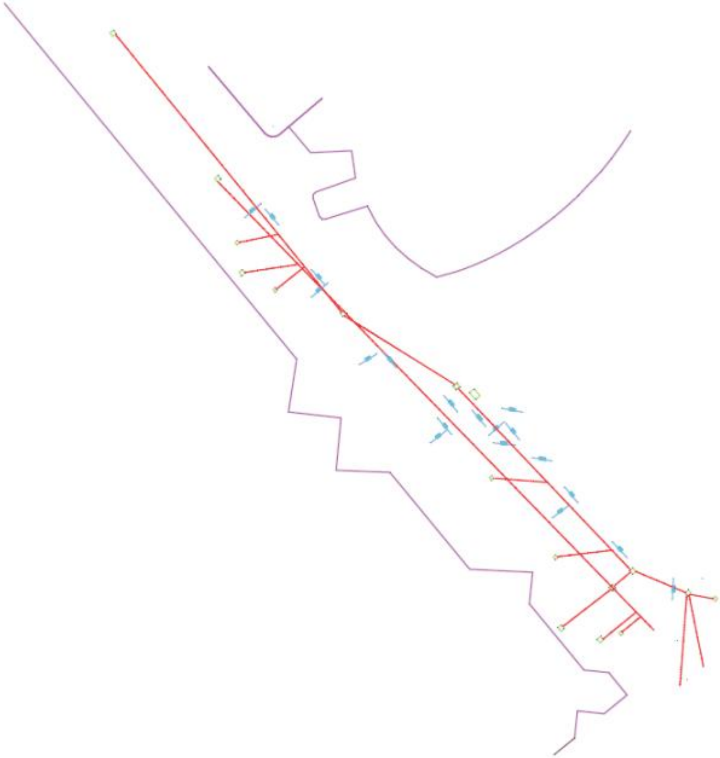
UNIVERSITY OF LEEDS



Experiments

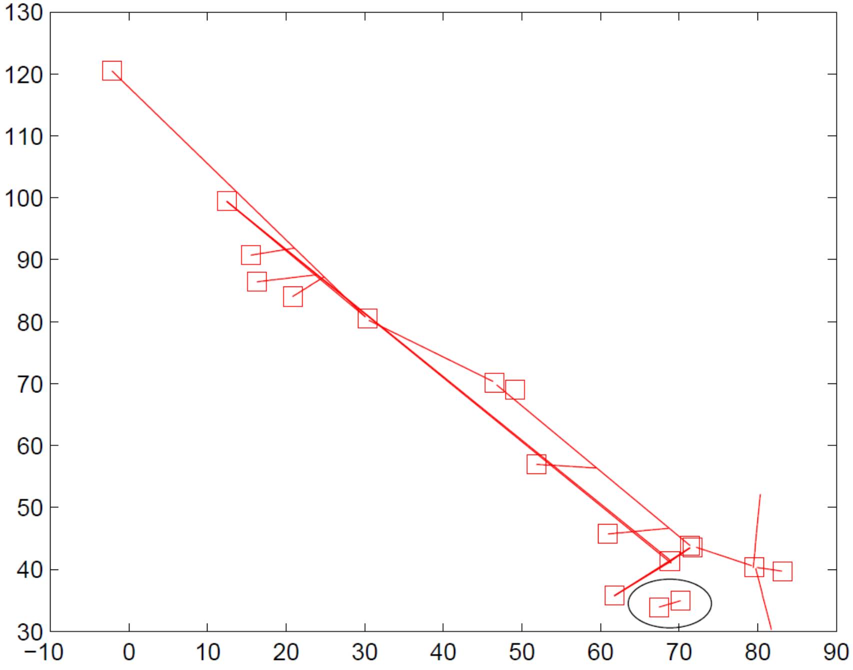
UNIVERSITY OF LEEDS

- Two real data sets from the UK: each data set consists of statutory records, set of GPR point scans and the street survey results.

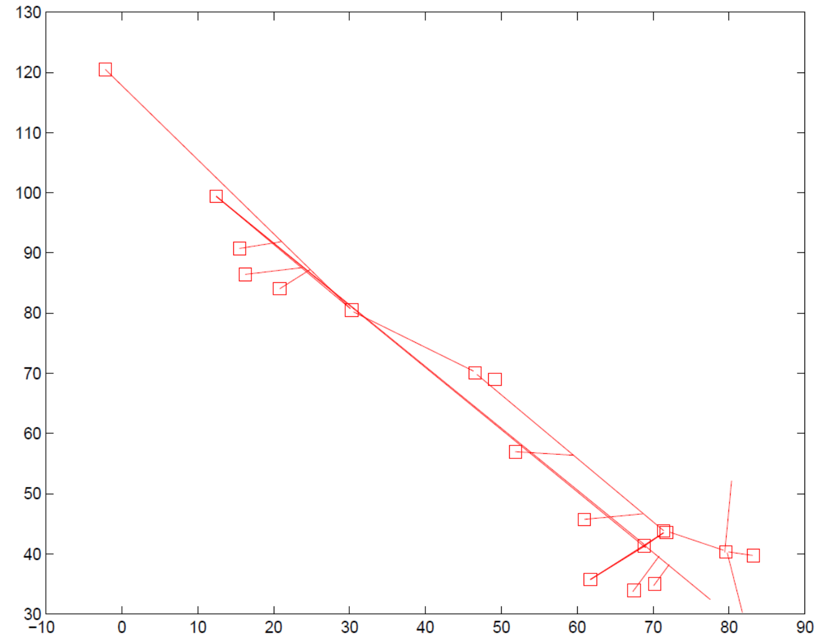


BDF Output

UNIVERSITY OF LEEDS



Site 1 without statutory record

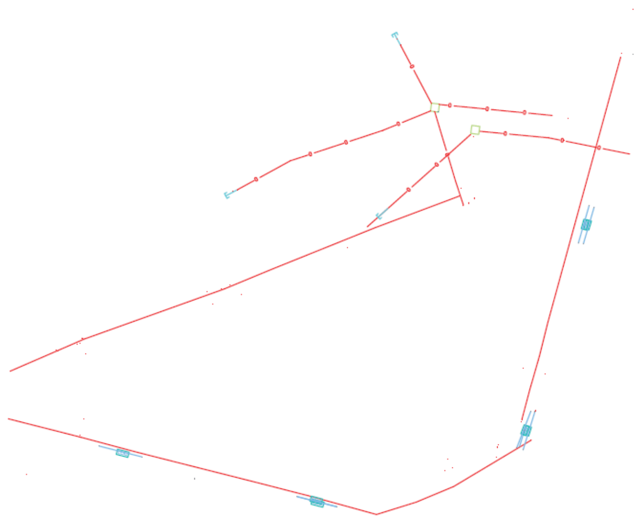


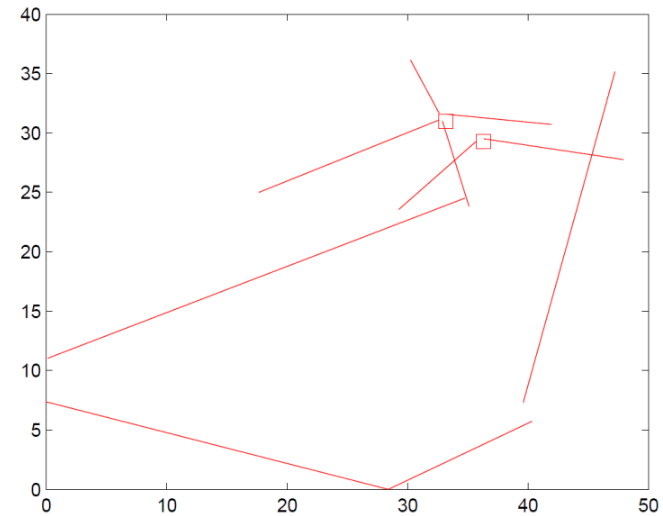
Site 1 with statutory record

□ - manhole, red line - pipe

Birmingham City University Water Pipes

UNIVERSITY OF LEEDS

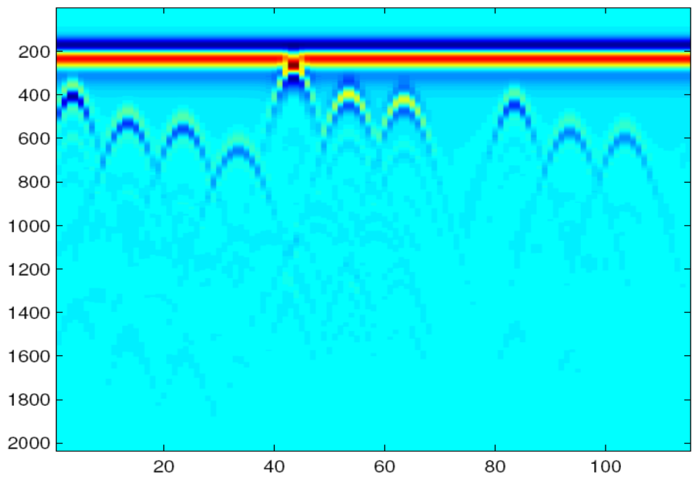




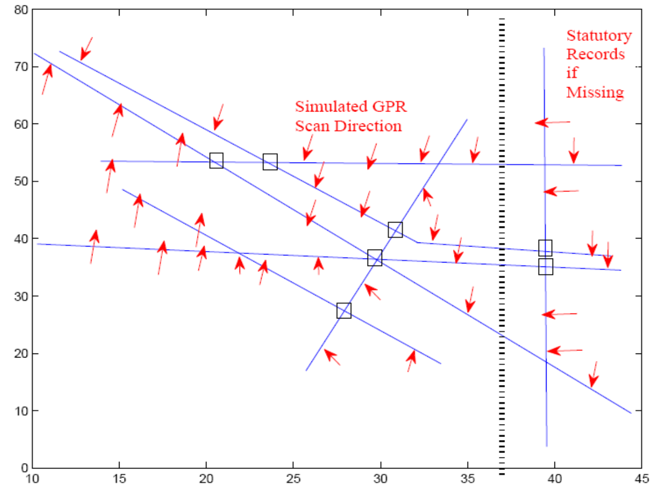
No connection errors for site 2 data.

Simulated Data

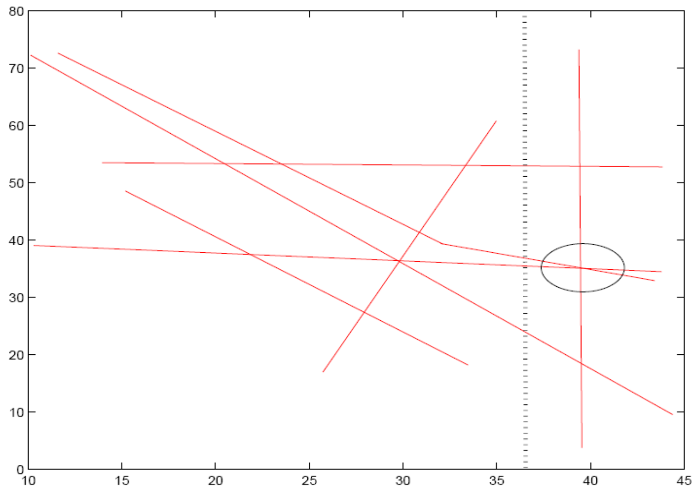
UNIVERSITY OF LEEDS



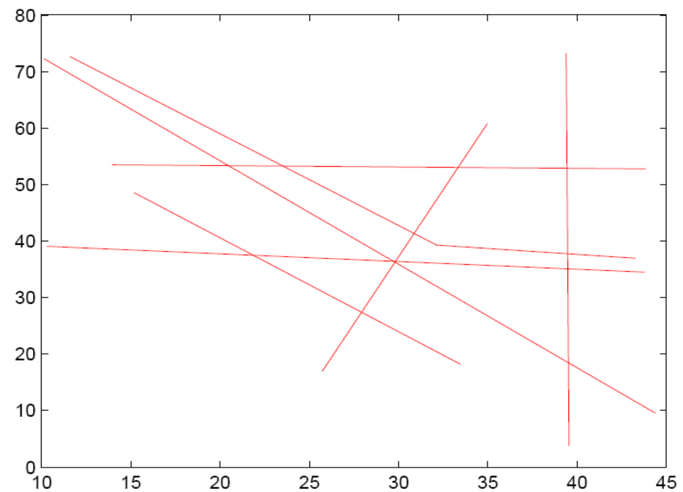
Simulated GPR Result



Simulated Data Set



Mapping without statutory record



Mapping with statutory record

Data Sets and Connection Errors

Data	Area (m×m)	# Pipe segments	# Manhole	#Scans
R1	100×100	19	18	23
R2	100×100	10	2	26
S1	100×100	22	7	36
S2	100×100	38	10	54
S3	100×100	44	18	76
S4	100×100	61	21	93

Summary information for two real data sets (R1, R2) and four simulated data sets (S1, S2, S3 and S4).

#error	BDF	BDF**V**	BDF**L**	BDF\\GPR	JCBB(**O**)	**V**	Time(s)
R1	0	2	1	3	5	1	6.6
R2	0	0	0	0	0	0	2.9
S1	0	3	2	5	5	2	9.3
S2	1	2	1	4	3	2	10.2
S3	1	1	1	6	4	1	13.7
S4	2	4	3	8	6	3	14.6

The connection errors of BDF, BDF without statutory records (BDF\V), BDF without pipe linearity assumption (BDF\L), BDF without GPR (BDF\GPR), JCBB on only GPR/manhole survey (JCBB(O)), and statutory records (V) on two real-world data sets (R1 and R2) and four simulated data sets (S1-S4).

Spatial Error or BDF Related Algorithms

UNIVERSITY OF LEEDS

spatial error	BDF	BDF\V	BDF\L	BDF\GPR	JCBB(O)	V
$E(x,\theta)_{R1}$	0.3,3.3	0.5,7.2	0.4,4.7	0.7,5.8	0.7,9.4	0.8,6.5
$C(x,\theta)_{R1}$	0.2,3.6	0.3,6.5	0.2,4.5	0.4,4.6	0.3,10.8	1,8
$E(x,\theta)_{R2}$	0.4,2.8	0.4,6.3	0.4,4.1	0.9,5.4	0.8,10.1	0.6,6.1
$C(x,\theta)_{R2}$	0.2,3.8	0.3,7.3	0.3,4.9	0.6,5.9	0.4,12.6	1,8
$E(x,\theta)_{S1}$	0.3,3.1	0.5,7.9	0.3,5.0	0.9,8.6	0.8,10.6	0.7,6.9
$C(x,\theta)_{S1}$	0.2,3.8	0.3,7.1	0.3,4.8	0.5,7.3	0.4,12.2	1,8
$E(x,\theta)_{S2}$	0.3,3.4	0.5,7.9	0.5,5.1	0.9,9.3	0.8,10.7	0.7,7.1
$C(x,\theta)_{S2}$	0.2,3.8	0.3,7.1	0.3,4.8	0.5,7.6	0.4,12.2	1,8
$E(x,\theta)_{S3}$	0.4,3.6	0.8,8.4	0.5,5.3	0.9,10.2	0.9,11.2	0.8,7.2
$C(x,\theta)_{S3}$	0.2,3.8	0.3,7.1	0.3,4.8	0.6,8.9	0.4,12.0	1,8
$E(x,\theta)_{S4}$	0.4,3.8	0.9,8.6	0.5,5.6	1.0,11.1	0.9,11.5	0.8,6.8
$C(x,\theta)_{S4}$	0.2,3.8	0.3,7.1	0.3,4.8	0.5,8.1	0.4,12.1	1,8

- The spatial errors and uncertainty. E(x, ϴ) represents the mean spatial distance (in metres) from the real PoI from the estimated PoI and the mean difference (in degrees) of real pipe direction and the estimated direction.
- C(x, ϴ) stands for the uncertainty of these two terms.

Summary of experiments

UNIVERSITY OF LEEDS

- BDF with full data sources outperforms other algorithms.
- JCBB with only GPR/manhole survey is the worst.
- BDF**V** is equal or inferior to BDF**L**, indicating the statutory records usually contain more information, if presented, than the pipe linearity assumption.
- In general it is very unlikely that only using observations of street furniture (such as manholes) and statutory records will give good results (some buried utilities may well have no such street furniture in the surveyed area). Therefore, BDF\GPR generates inferior performance than BDF with all data sources.
- The computation time are recorded on a 2.4Ghz laptop with 4GB memory on a single core. Clearly, the algorithm can operate in real time as data is gathered (given the push speed of the GPR).

All these results confirm the benefits of inclusion of more data sources and the effectiveness of the BDF algorithm in utility mapping.

Conclusion and Future Work

UNIVERSITY OF LEEDS

- Previous approaches to produce buried utility pipeline maps depend on manual drawing and expert interpretation of GPR scans.
- Our work represents the first attempt to automatically map utility data from sensor input.
- Algorithms for Bayesian data fusion (BDF) of multiply data sources to connect these manholes and GPR scan locations.
- Comparison of BDF methods with different combinations of data sources.
- The uncertainty on the location and direction of pipes are both considered in the algorithm.

Future Work

- Incorporate expectations from other sensors, e.g. sonar, passive electromagnetic and low frequency electromagnetic.
- Real-time on-line operation (cf SLAM)
- **RF VACANCY**

References

H. Chen and A. G. Cohn. Probabilistic conic mixture model and its applications to mining spatial ground penetrating radar data. In Workshops of SIAM Conference on Data Mining (WSDM10), 2010.

H. Chen and A. G. Cohn. Probabilistic Robust Hyperbola Mixture Model for Interpreting Ground Penetrating Radar Data. In 2010 IEEE World Congress on Computational Intelligence. Barcelona, July, 2010.

Chen, H. and Cohn, A.G. Buried Utility Pipeline Mapping based on Street Survey and Ground Penetrating Radar, Proceedings of the European Conference on Artificial Intelligence (ECAI'10), Lisbon, Portugal, 2010.