

UK Rail Data from a Technical Perspective

Peter Hicks
OpenTech 2011

About me

- Network and Telecoms Engineer
- 15 years experience as a commuter
- Supporter of rail travel
- Conversant with transport technology
- Keen cyclist (spot the irony)
- Foil fencer

Transport Projects

- TubeHorus
 - Interface to TfL's Tracknet data
 - Version 2 planned
 - <http://www.transporthacker.com/tubehorus/>
- TransportHacker
 - Mash-up of Highways Agency and TfL Streets data
 - Very 'beta'
 - <http://www-staging.transporthacker.com/>

Agenda

- Timetables in CIF
- Real Time Data
- TSDBExplorer
- Where do we go from here...?

Agenda

- Timetables in CIF
- Real Time Data
- TSDBExplorer
- Where do we go from here...?

Timetables in CIF

Where do CIF files come from?

- Network Rail's TSDB - Train Service Database
- CIF produces extracts
 - Filter by location - "Only trains through Watford Junction"
 - Filter by train type - "Only passenger trains"
 - Filter by features - "Only reserveable trains"
- Schedules contains 'runs from' and 'runs to' days, as well as 'runs on days'
- Exception records override or cancel data

Timetables in CIF

CIF file format

- “The file is sequential containing fixed length 80 character records. It will contain different record types which can be identified by the 'record identity', the first two bytes of a record. Individual records can be updated or deleted in CIF updates”
- Around 450Mb for a ‘full extract’
- Nightly updates - ‘update extract’ - usually small
- Search for ‘network rail common interface file’ for the specification

Timetables in CIF

What's in a CIF extract?

1. A set of locations (TIPLoCs)

Timetables in CIF

What are TIPLOCs?

- Timing Information Point LOcation
- TIPLOCs used for train schedules
- Nearly everything has a TIPLOC
 - Stations
 - Parts of a station (e.g. Reading 4A+4B, Reading)
 - Sidings
 - Signals
 - Depots
- CRS and NLC codes used for ticketing

Timetables in CIF

An example of a TIPLOC

Field	Value
TIPLOC Code	EUSTON
NALCO (National Location Code)	144400
TPS Description	LONDON EUSTON
STANOX	72410
CRS Code	EUS
Description	LONDON EUSTON

...but no geography information!

Timetables in CIF

What's in a CIF extract?

1. A set of locations (TIPLOCs)
2. A set of associations between schedules

What's an association?

- A link between two schedules
 - Train X forms Train Y
 - Train X splits to form Train Y
 - Train X joins with Train Y
- Useful for real-time reporting
 - “If Train X is 15 minutes late at its destination and the train it forms departs 10 minutes later, then that train will probably be 5 minutes late departing”

Timetables in CIF

What's in a CIF extract?

1. A set of locations (TIPLOCs)
2. A set of associations between schedules
3. A set of schedules

Timetables in CIF

What makes up a schedule?

- A 'Basic Schedule' record
 - Train identity (e.g. 2J34)
 - Train class (e.g. Express Passenger)
 - Timed speed (e.g. 90 mph)
 - Catering (Trolley, Buffet, Restaurant, None)
 - Sleeper service
- A 'Basic Schedule Extended' record
 - Operator
- A 'Location Origin' record
- Zero or more 'Intermediate Location' records
- A 'Location Terminate' record

Timetables in CIF

What's in a location record?

- Arrival and Departure times
- Public Arrival and Departure times
- Passing times
- Platform information
- Engineering and performance allowances
- Activities

Timetables in CIF

A sample schedule

UID L90416

Train 2J34

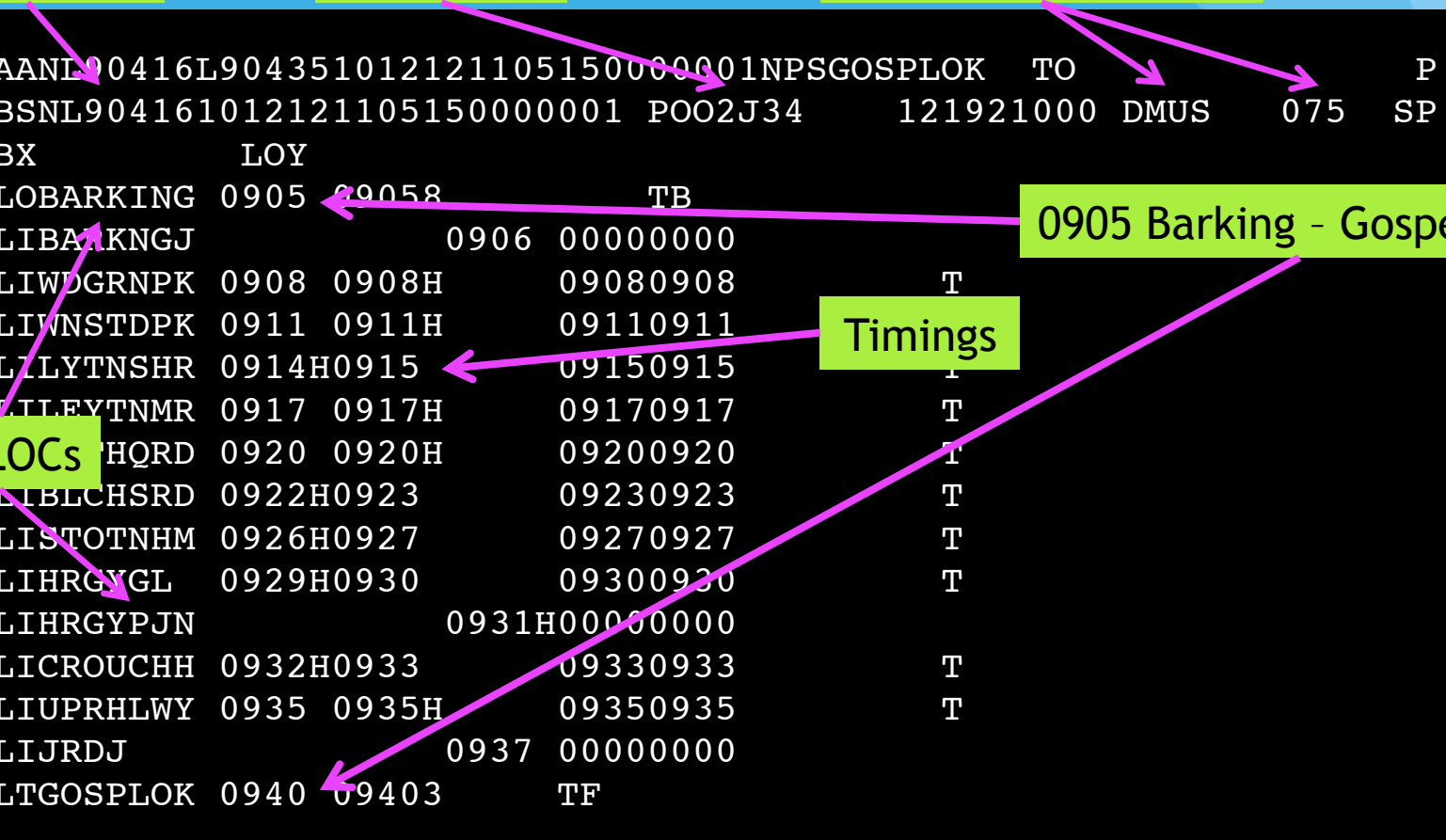
Diesel train, 75mph

AANL	0416	L904351012121105150000001	NPSGOSPLOK	TO		P
BSNL	904161012121105150000001	POO2J34	121921000	DMUS	075	SP
BX	LOY					
LOBARKING	0905	09058	TB			
LIBARKNGJ		0906	00000000			
LIWGRNPK	0908	0908H	09080908			T
LIWNSTDPK	0911	0911H	09110911			
LILYTNshr	0914H	0915	09150915			T
LILEYTNMR	0917	0917H	09170917			T
LIPLOCS	0920	0920H	09200920			T
LIBLCHSRD	0922H	0923	09230923			T
LITOTNHM	0926H	0927	09270927			T
LIHRGYGL	0929H	0930	09300930			T
LIHRGYPJN		0931H	00000000			
LICROUCHH	0932H	0933	09330933			T
LIUPRHLWY	0935	0935H	09350935			T
LIJRDJ		0937	00000000			
LTGOSPLOK	0940	09403	TF			

0905 Barking - Gospel Oak

Timings

TIPLOCs



Timetables in CIF

What can we do with it?

- 'Build Your Timetable' service
- Minimal single-leg journey planner
- Data visualisation
- Service comparison
- Extend with real-time running data

Agenda

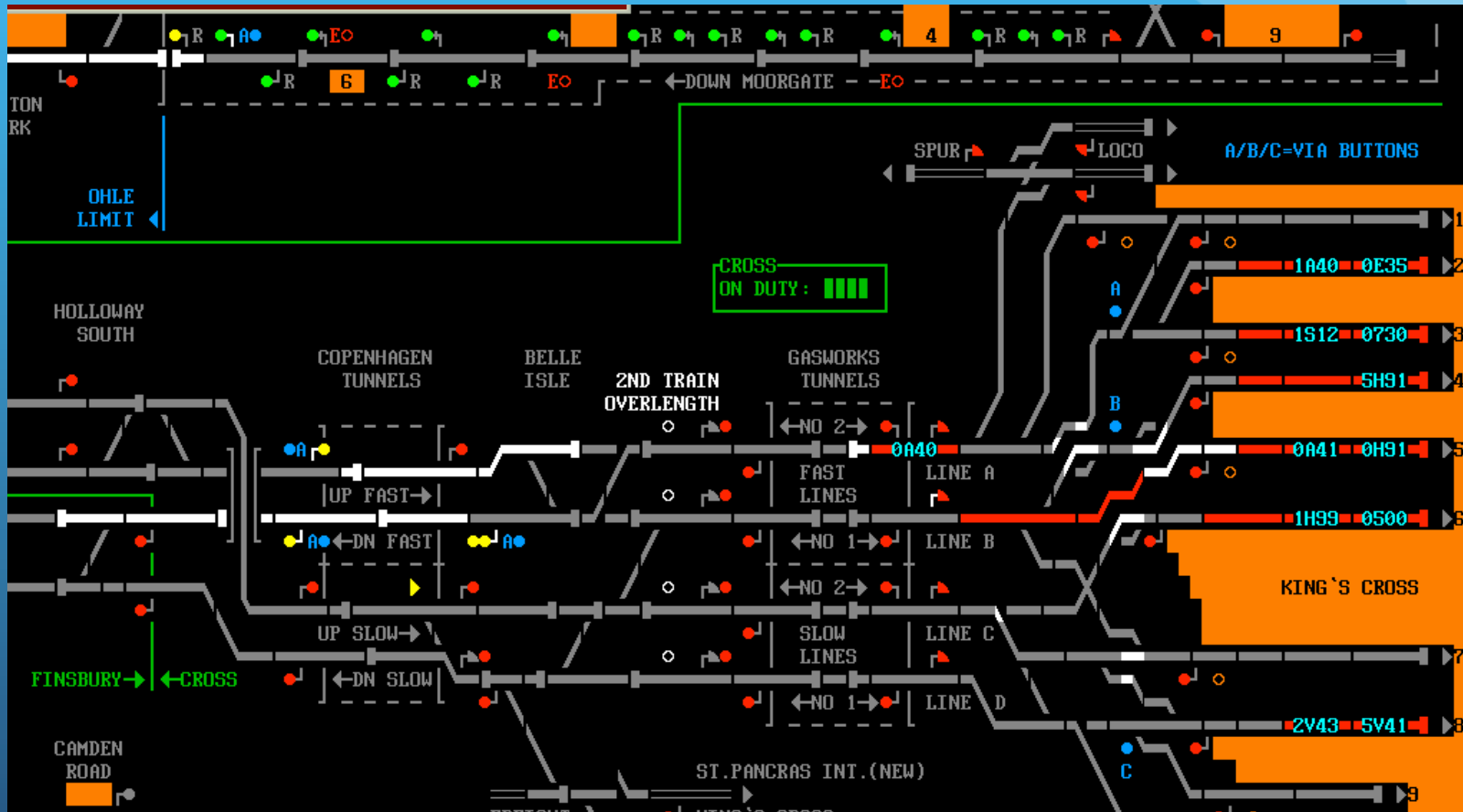
- Timetables in CIF
- **Real Time Data**
- TSDBExplorer
- Where do we go from here...?

How do you identify a train?

- All trains are identified by a Train ID (e.g. 2J34)
 - Not used for public-facing information
 - Unique within a signalling area within a 6h period
- Trains in CIF have a UID - not used operationally
- Signallers know where every train in their area is

Real Time Data

Where are my trains?



Screenshot from SimSig Kings Cross

Where are my trains?

- Track Circuit/Axle Counter operation
 - Automatic and very accurate
 - Sections can be very long
 - Impossible to tell if a train has stopped in a section
- Mass Detectors
 - “A train has passed over me”
 - Used for CIS, not always linked to signalling systems
 - “The train now approaching platform 4...”
- Manual input

How do you get the information?

- From ATOC's DARWIN system
 - If you meet their self-written criteria
 - If you pay what they want you to pay
- From LiveDepartureBoards.co.uk
 - Screen-scrape!
 - ...except this is against the AUP
- TRUST - Train Running System TOPS
 - IBM 3270-based
 - No public access
 - Only usable by 'experts'

How do you get the information?

- From Network Rail's TD.net system...?
 - “TD.net is a “publish and subscribe” architecture designed to enable the publication of train-related data internally within Network Rail and externally with other industry parties. Subscribing clients include system integrators, innovation partners, and train operators”[1]

Real Time Data

What does TD.net give?

- Structured XML messages
- Single interface to positioning data
- Train Describer messages
- Train positioning and movement event data
- Very Short Term Plan (VSTP) schedules
- Temporary Speed Restriction data
- TRUST incident and delay messages

Agenda

- Timetables in CIF
- Real Time Data
- **TSDBExplorer**
- Where do we go from here...?

TSDBExplorer

What is it?

- Proof of Concept
- Work-in-Progress
- Ruby on Rails
- CIF processing engine
- HTML-based query front-end
- Very 'alpha'

What does it look like?

TSDB Explorer

Schedule for train 2J34 (C40482), valid MTWThF from Mon 13 Dec 2010 to Fri 20 May 2011

Location					Times		Allowances		
					Public times in red text				
Name	Platform	Line	Path	Activity	Arrival	Departure	Eng	Pth	Pfm
Dorridge	3	UDP	-	TB		1128	1128		
Bentley Heath Crossing						<i>Pass 1130</i>			
Widney Manor				T	1132	1132	1132H	1132	
Solihull				T	1135	1135	1136	1136	
Olton				T	1139	1139	1139H	1139	
Acocks Green				T	1141H	1142	1142	1142	
Tyseley	2	FL				<i>Pass 1143H</i>			
Small Heath South Junction		SH				<i>Pass 1144H</i>			
Birmingham Moor Street	2			T	1147H	1148	1148H	1148	
Birmingham Snow Hill	1			T	1150H	1151	1153	1153	
Jewellery Quarter				T	1155	1155	1155H	1155	
The Hawthorns				T	1159H	1200	1200	1200	
Smethwick Galton Bdg H.I.				T	1202	1202	1203	1203	
Smethwick Junction						<i>Pass 1203H</i>			
Rowley Regis				T	1207H	1208	1208H	1208	
Cradley Heath				T	1213	1213	1213H	1213	1
Stourbridge Junction		-		TF	1219	1219	-	-	-

TSDBExplorer

How do I get it?

- Git repository - <http://git.poggs.com/>
- You will need CIF data!
 - I have it
 - I am not sure if I can distribute it...
 - ...but I am working to produce a set of sample data
 - You might be able to get it from Network Rail
 - Ask me, don't flood Network Rail with queries!

TSDBExplorer

What's planned?

- Adding real-time data - 'Open Rail Data'
- Produce CIF data in other formats
- RESTful and XML-based APIs
- ...and things we haven't even thought about!

Agenda

- Timetables in CIF
- Real Time Data
- TSDBExplorer
- Where do we go from here...?

Where do we go from here...?

Fares Information

- Data owned by ATOC
- Fares Manuals
- Already available electronically
- Avantix Traveler CD-ROM available from TSO
- Not open data... yet
- Commercially sensitive?

Where do we go from here...?

Routeing Guide

- Complicated model of ‘permitted routes’ for an ‘Any Permitted’ routed ticket
- Many exceptions, called ‘easements’
- Data owned by ATOC
- Already available electronically
- Not open data... yet

Where do we go from here...?

Geospatial Data

- Topology in electronic format
 - Draw maps, show where trains are?
 - Compare road and rail disruption
 - OpenStreetMap has good detail
- Open access to ATOC's station information
 - It's there to help the public!
 - Probably not open data... yet

Where do we go from here...?

Access to TD.net

- We're working on it

Where do we go from here?

Talk, talk, talk

- Talk to us and the community
- What do you want to do with rail data?
- Show your ‘proof of concept’ - get people excited
- Show TOCs, ATOC and Network Rail there is demand for Open Rail Data

Get in touch!

- Email: peter.hicks@poggs.co.uk
- Twitter: @poggs
- Blog: <http://blog.poggs.com/>



Open Discussion