
Reliable data collection in highly disconnected environments using mobile phones

Brian DeRenzi

Department of Computer Science & Engineering
University of Washington

Joint work with Yaw Anokwa, Tapan Parikh, and Gaetano Borriello

Data collection

- n Four and a half billion people in the developing world
 - q All need access to basic services like healthcare

- n Offering and maintaining services requires information
 - q Digitizing information has advantages, but how do we transport that data?

Data transmission

- n Most rural locations have limited or no network coverage
 - q In rural southern Tanzania, only 34% (n=134) of health facilities have mobile coverage¹

- n Use data mules to transmit data
 - q Mobile devices that ferry data across regions of little or no connectivity

¹ Lindi & Mtwara Health Facility Survey 2004 IPTi in southern Tanzania IHDR



“Some details”

- n Collection of medical data
 - q Obtained through examinations and interviews
 - q Often collected in isolated rural villages
- n Difficulties
 - q Expensive to collect (field workers)
 - q High error rates (misinterpretations, incomplete data, incomprehensible notes, etc.)
 - q High-latency to usefulness (delay in data entry)

Data Collection Days



Village to Village

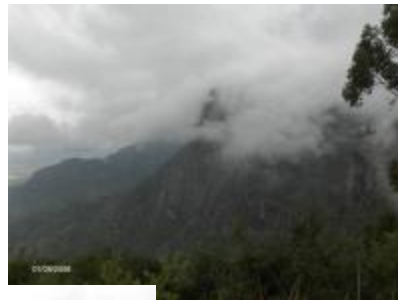


27 August 2007

NSDR 2007 - Kyoto, Japan

7

Rural Tanzania



27 August 2007

NSDR 2007 - Kyoto, Japan

8

Data Collectors



27 August 2007

NSDR 2007 - Kyoto, Japan

9

Objectives

- n Simplify data collection
 - q Dynamic forms on mobile device
 - q Easy to add metadata (pictures, audio, video)
 - q IDs, biometrics, geography for identification
 - q Low-latency collection
- n Safeguard data
 - q Automatic backup of data
 - q Get it to servers as soon as possible

27 August 2007

NSDR 2007 - Kyoto, Japan

10

Outline

- n Motivation
- n System components
- n New applications
- n Current progress
- n Next steps
- n Acknowledgements

Definitions

- n Nodes
 - q Mobile devices out in the field
- n Neighbors
 - q Nodes are neighbors if their devices can talk to one another
 - q One-hop neighbors
- n Server/Safe store
 - q A centralized server where data is safe
- n Source
 - q A node which is sending, or has sent, data to a server

System components

n Data Routing

- q Use existing social networks to route data in disconnected environments

n Location Profile

- q Using GPS traces, build a profile of where the person has gone and where (s)he might go next
- q Automatic exchange of location profile to determine good candidates to ferry data

System components continued

n Receipt acknowledgements

- q Due to collection cost, do not delete data until it is been marked as received
- q Path back to source will not necessarily be the same

n Storage reclamation

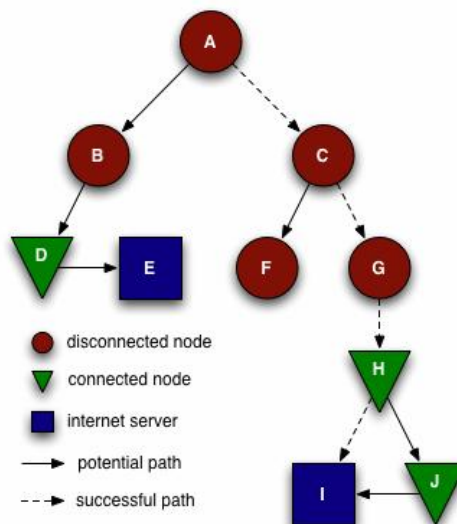
- q Data furthest from the source is deleted first
- q Devices with data one hop away from source explicitly notify if there is a need to remove the data
- q Source notifies those one hop away when a receipt is received

Routing to the server

- n Forward to neighbors
 - q Increment hop count
- n 1-hop neighbors are special
 - q Most likely to be of same organization
 - q Backup data for each other
- n Multi-hop neighbors are opportunistic best effort
 - q Preference for large mules
 - q Busses, trucks, government vehicles

From field worker to server

- n Not all paths will be successful

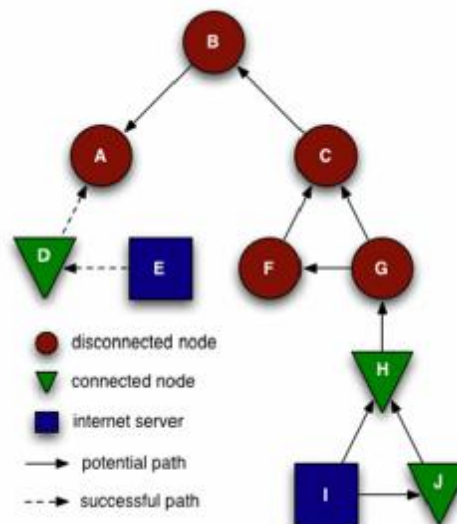


Routing from server

- n New data and acks of delivered data
- n Acks
 - q Mule that delivers gets ack
 - q But also sent along other routes
- n Data route determined by location profiles
 - q Collected by devices GPS unit
 - q Routed to mules likely to be heading where the packet has to go (may not be same place as teams move, multiple possible geographic destinations)
 - q Also based on social links

From server to field worker (node)

- n Nodes are mobile and may have moved, so the path will not necessarily be the same.



Safeguarding Data

- n Share data among members of medical team
 - q Groups of field workers travel together
 - q Backup each others' collected data
- n Leverage social/work movements to move data
 - q Each field worker can “mule” data to where there is connectivity
 - q Exploit encounters with other NGO workers, truck drivers, government workers, etc.

Other considerations

- n Data Prioritization
 - q Send only core data first
 - q Send smallest data first
- n Usage Incentives
 - q Could provide airtime for those transferring data
- n User Privacy
 - q PKI to protect patient data being transmitted
 - q Variable granularity for location profiles

Outline

- n Motivation
- n System components
- n **New applications**
- n Current progress
- n Next steps
- n Acknowledgements

Applications

- n Data collection for surveys in rural villages
 - q Public health maternal and neonatal survey for the IPTi project¹
 - q Child morbidity monitoring project¹
 - q Stillbirth and miscarriage survey²
 - q Accompagnateurs in PIH
- n Drug monitoring and patient data
 - q Dispensaries refer to health clinics that refer to district hospitals
 - q Data from medical protocols on mobile devices

¹ Ifakara Health Research and Development Centre

² Rachel Haws, Johns Hopkins University

Medical Protocols on mobile devices

- n Protocols exist to aid clinicians
- n Currently delivered on paper-based chartbooks.
- n Using digital devices has a number of advantages
 - q Patient record
 - q Updatable protocols
- n Need to get data back to policy makers



Outline

- n Motivation
- n System components
- n New applications
- n Current progress
- n Next steps
- n Acknowledgements

Regions of Operation

- n Summer 2007
 - q Brian DeRenzi with D-Tree in Mtwara, Tanzania
 - q Yaw Anokwa with PIH in Rwigyema, Rwanda
- n Summer 2006
 - q Brian was in Kilimanjaro and Tanga regions in Tanzania



Status

- n Composed pieces on Nokia phones
 - q First version of dynamic forms for data
 - q GPS data collection (separate unit)
 - q Message passing from device to device
- n Prototyped simple email application
 - q Computer to phone via MSR's smstoolkit
 - q Phone to phone via Bluetooth

Other Applications

n Guatemala

- q Remote village with cell phone connection only at top of nearby mountain
- q Local entrepreneur shuttles messages for villagers (on paper)
- q Could support voice, SMS, and e-mail by muling from villagers' own phones
- q Better privacy, shared use of the only phone in the village



Outline

- n Motivation
- n System components
- n New applications
- n Current progress
- n Next steps
- n Acknowledgements

Next steps

- n Development of the system over the next year
 - q Add more automation to current prototypes
- n Deployment and field testing next summer and the following year
 - q Explore the open questions about requirements

Acknowledgments

- n Thank you
 - q SIGCOMM/NSDR
 - q International Service Learning (ISL)
 - q Ifakara Health Research & Development Centre (IHRDC)
- n Questions/comments?

