

Ocupado: Visual Analytics for Occupancy Applications

Tamara Munzner
 Department of Computer Science
 University of British Columbia

THE UNIVERSITY OF BRITISH COLUMBIA
 DESIGNING for PEOPLE

Cisco Toronto
 26 Jun 2018

www.cs.ubc.ca/~tmm/talks.html#cisco18 @tamaramunzner

Ocupado project

- occupancy data for facilities management
 - estimate human occupancy of buildings using mobile device connections as common denominator
 - innovative uses for CMX data
 - create visual and predictive decision-support tools
 - visual analytics interface to make data actionable by people
 - investigate multiple stakeholder contexts of use
 - from energy management to space planning and beyond

www.cs.ubc.ca/~tmm/talks.html#cisco18

Ocupado collaboration: Partners

- visualization and data analysis: UBC Computer Science
 - led by Tamara Munzner
 - students: Michael Oppermann, Yann Dubois
- building management systems and data recording: Sensible Building Science
 - UBC Sustainability spinoff
 - led by Stefan Storey
- machine learning: UBC Statistics
 - led by Jeff Andrews
- networking infrastructure: Cisco
 - liaison: Rob Barton

www.cs.ubc.ca/~tmm/talks.html#cisco18

Ocupado collaboration: Funding

- kickstarted by Cisco funding (\$25K)
 - *Locational Service Analytics: Machine Learning and Data Visualization for CMX Data Applications*
- matched 3.5x
 - UBC Campus as a Living Lab (\$41K)
 - NSERC Engage (\$25K)
 - planned: MITACS (\$25K)
- substantial increase in project scope & duration
 - 40% spent over 16 months from May 2017 - Sep 2018
 - we're now at month 13
 - 60% to spend in 24 months from Oct 2018 - Aug 2020

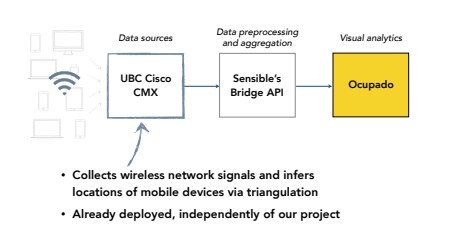
www.cs.ubc.ca/~tmm/talks.html#cisco18

Project threads: Completed to date

- visualization research
 - requirements analysis
 - visualization prototyping in Sandbox environment
 - experiment w/ static data
 - integrate with live data
- SBS Bridge2 product
 - develop & deploy
 - integrate basic ML
- machine learning research
 - basic prediction: short & long-term forecasting

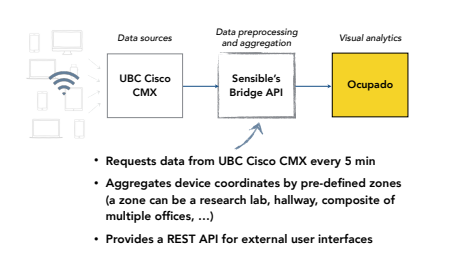
www.cs.ubc.ca/~tmm/talks.html#cisco18

Data architecture



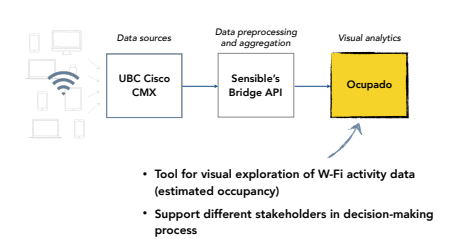
www.cs.ubc.ca/~tmm/talks.html#cisco18

Data architecture



www.cs.ubc.ca/~tmm/talks.html#cisco18

Data architecture



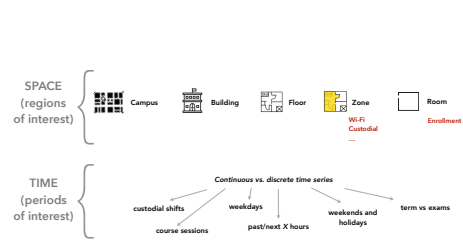
www.cs.ubc.ca/~tmm/talks.html#cisco18

Data: Wifi as proxy for human occupancy

- wifi device activity strongly correlated with occupancy
 - rough proxy for headcounts in rooms
 - device counts every 5 minutes, per zone
 - good spatial precision if zone large enough
 - rooms with multiple people, not single-person offices
 - excellent temporal resolution
- privacy preserving architecture
 - keep only counts per zone per time slice
 - no tracking of individuals or trajectories
 - privacy built in to SBS Bridge infrastructure at fundamental level
 - MAC addresses thrown away, not stored
 - we'd love CMX protocol change so they're not sent out!
 - (3 month delay in data gathering due to UBC Legal concerns)

www.cs.ubc.ca/~tmm/talks.html#cisco18

Data abstraction



www.cs.ubc.ca/~tmm/talks.html#cisco18

Task analysis: Facilities management stakeholders

- known in advance
 - energy systems
 - SBS first product: occupancy for HVAC control
- identified as high priority
 - space planning
 - informal learning spaces
 - classroom services
 - custodial services
 - building managers
- investigated and considered lower priority
 - risk management
 - security and parking
 - transportation

www.cs.ubc.ca/~tmm/talks.html#cisco18

Task analysis: Example stakeholder questions

- only basic query handled by previous SBS interface
 - What is the current activity level of a specific region?
- many stakeholder questions require bigger picture
 - Which regions are busy/quiet now?
 - Which regions were heavily used and are empty now?
 - What does the long-term activity profile of region X look like?
 - What is the typical usage pattern of a specific region?
 - weekdays vs weekends/evening/holidays, according to shift boundaries
 - How does the utilization differ between regions?
 - for subset based on size, space type or other attributes
 - What is the predicted activity for a region in the next X hours?
 - Which regions are normally heavy used but quiet now? (or vice versa)
 - detecting current anomalies vs. average patterns

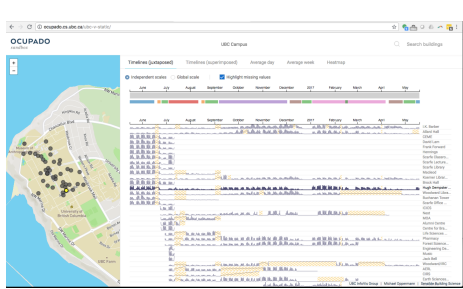
www.cs.ubc.ca/~tmm/talks.html#cisco18

Ocupado Sandbox

- alpha 0.1
- Dynamic filtering, slicing, and sorting of regions

www.cs.ubc.ca/~tmm/talks.html#cisco18

Static data: Test deployment, obvious gaps



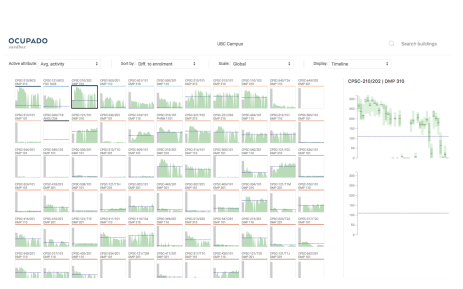
www.cs.ubc.ca/~tmm/talks.html#cisco18

Ocupado Sandbox

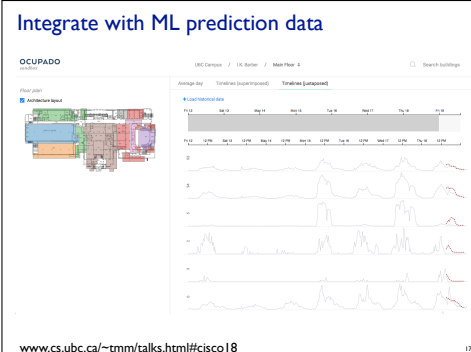
- alpha 0.2
- Integrate static data with other data sources: course schedules, predictions

www.cs.ubc.ca/~tmm/talks.html#cisco18

Scheduling data: Actual vs enrolled in courses



www.cs.ubc.ca/~tmm/talks.html#cisco18

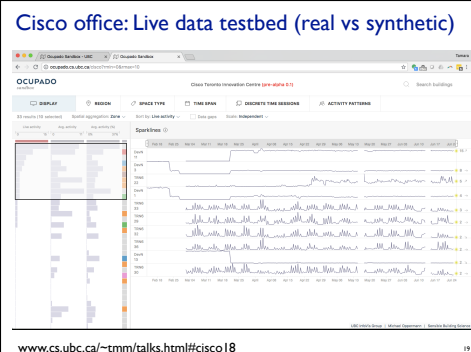


Ocupado Sandbox

alpha 0.3

- Flexible visual exploration interface between the user and the Bridge API
- Integration of live activity data
- Presets for quickly answering common domain questions
- URL bookmarks for replicating and sharing a certain application state

www.cs.ubc.ca/~tmm/talks.html#ciscol8



Ocupado Sandbox

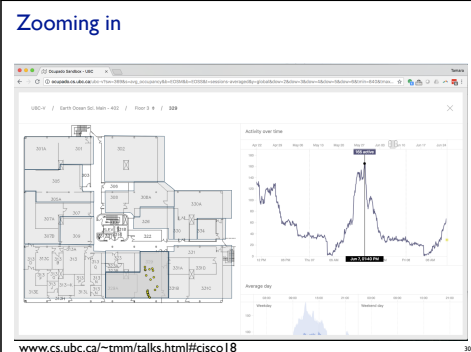
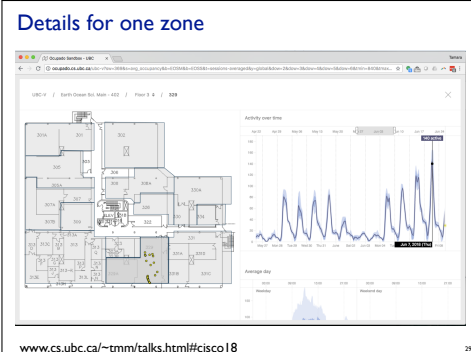
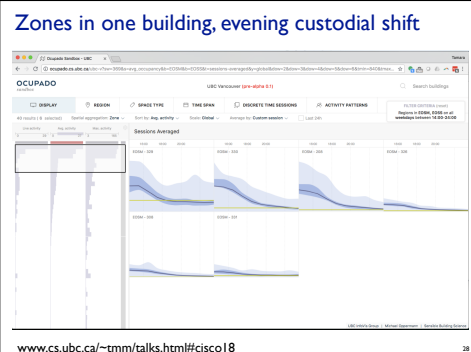
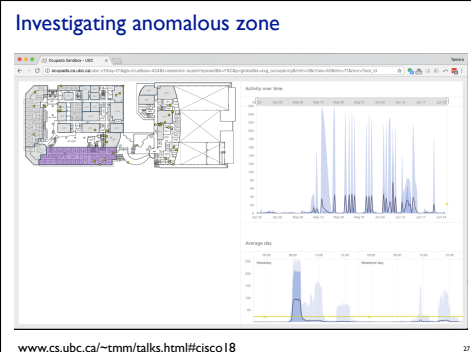
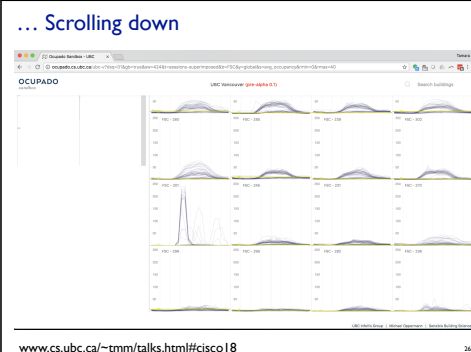
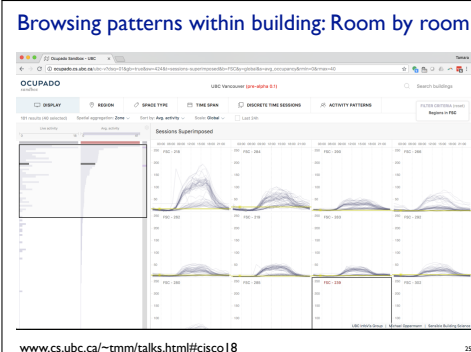
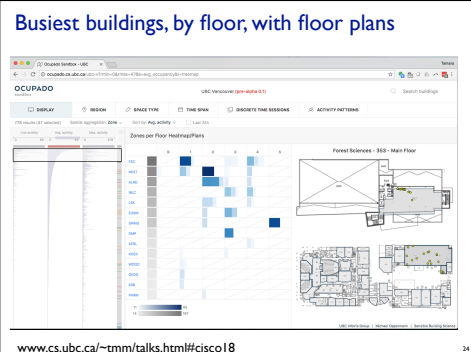
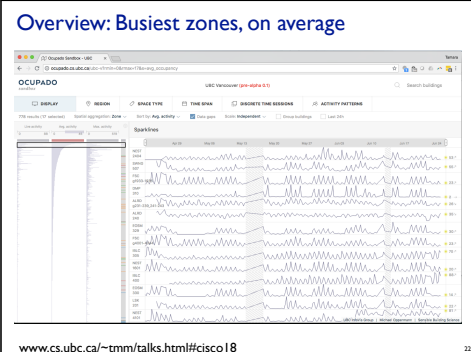
alpha 0.4

- Live data flowing from UBC
- Continued development of visual interface

www.cs.ubc.ca/~tmm/talks.html#ciscol8

Demo

www.cs.ubc.ca/~tmm/talks.html#ciscol8



Ocupado timeline: Milestones to Sep 2018

- May 2017 - Sep 2018**
 - UBC: Machine learning (prediction), Task/requirements analysis
 - Yann Dubois (BS/Intern), Munzner, Andrews
 - May - Sep 2017
- SBS: Bridge2 (data infrastructure), Machine learning (integration)**
 - Felipe Deo, Nick Bradley (MSc/Intern)
 - May 2017 - Apr 2018
- UBC: Visualization prototypes, Task/requirement analysis**
 - Michael Oppermann (PhD), Munzner
 - Aug 2017 - Sep 2018

Final 1.3 yrs: 40% funds used

- May - Dec 2017: 18K (Cisco/CLL)
- (SBS)
- Jan - Sep 2018: 25K (NSERC Engage)

Final 2 yrs: 60% funds left

- Sep 2018 - Aug 2020: 48K (Cisco/CLL) + 29K (planned MITACS)

www.cs.ubc.ca/~tmm/talks.html#ciscol8

Project threads

- visualization research
 - requirements analysis
 - visualization prototyping in Sandbox environment
 - experiment w/ static data
 - integrate with live data
 - customized visualization Skins for stakeholders
 - initial development (Jun 2018)
 - deployment & testing (Sep 2018)
 - exploit advanced ML in visualization (Sep 2019)
- machine learning research
 - basic prediction: short & long-term forecasting
 - semi-supervised asset tagging (Sep 2018)
 - advanced prediction: gaps & assets (Apr 2019)
- SBS Bridge2 product
 - develop & deploy
 - integrate basic ML
 - integrate advanced ML (Jan 2020)

www.cs.ubc.ca/~tmm/talks.html#ciscol8

Intellectual property

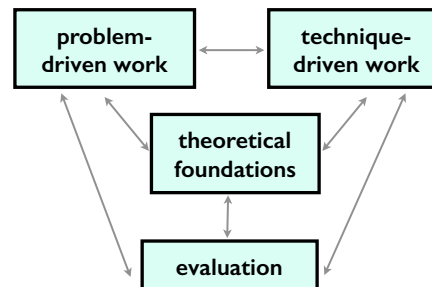
- open-source everything created at UBC
 - after moderate delay, under commercializable license
- unified whole
 - each part builds on and depends on others
 - impossible to disentangle IP into multiple buckets based on chronology/source
- benefits to partner companies
 - open-source Ocupado specifically designed as front end that interoperates with proprietary Bridge infrastructure from SBS
 - Ocupado+Bridge combination showcases benefits of occupancy tracking via Cisco CMX product
 - intellectual contribution of task analysis of stakeholder needs for different verticals is crucial but not patentable

www.cs.ubc.ca/~tmm/talks.html#cisco18

UBC InfoVis Group: Research Approach

www.cs.ubc.ca/~tmm/talks.html#cisco18

Research agenda: Interleaved angles of attack



www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven work

- design studies
 - in collaboration with target users
 - real data, real tasks
 - intensive requirements analysis
 - iterative refinement
 - deploy tools/systems
 - typical evaluation: field studies
- my strategy: opportunistic collaboration
 - many domains
 - both industrial and academic partners

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Tech industry

Heidi Lam

Diane Tang (Google)

SessionViewer: web log analysis
<https://youtu.be/T4MaTZd56G4>

Peter McLachlan

Stephen North (AT&T Research)

LiveRAC: systems time-series logs
<https://youtu.be/lDc3H0VSKw>

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Genomics

Aaron Barsky

Jenn Gardy (UBC Micro)

Robert Kincaid (Agilent)

Cerebral
<https://youtu.be/76HhG1FOngI>

Miriah Meyer

Hanspeter Pfister (Harvard)

MizBee

MulteeSum, Pathline

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Genomics

Joel Perstay

Cydney Nielsen (BC Cancer)

Variant View
https://youtu.be/AhDnv_qMXxQ

Ana Crisan

Jenn Gardy (UBC Public Health & BC CDC)

current work: genomic epidemiology

Zipeng Liu

current work: gene trees (UBC Zoology)

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Automotive, journalism

ReIX (BMW)
<https://youtu.be/89f9QXc6A04>

Michael Sedlmair

Jonathan Stray (Assoc Press)

Overview
<https://vimeo.com/71483614>

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Building mgmt, fisheries

Energy Manager

Matt Brehmer

Kevin Tate (Pulse/EnerNOC)

Vismon
<https://youtu.be/hkHtoS4VYmk>

Maryam Booshehrian

Torsten Moeller (SFU)

www.cs.ubc.ca/~tmm/talks.html#cisco18

Problem-driven: Current data science

Kimberly Dextras-Romagnino

current work: Segmentifier (Mobyfiy)

e-commerce clickstreams

build tools for human-in-the-loop visual data analysis

Michael Oppermann

current work: Ocupado (Sensible Building Science, Cisco)

wifi proxy for real-time building occupancy

integrate visual analytics and predictive ML for facilities management

www.cs.ubc.ca/~tmm/talks.html#cisco18

Technique-driven work

- scalable algorithms & systems
 - typical evaluation: computational benchmarks
- new layout & interaction techniques
 - typical evaluation: controlled experiments on human subjects

www.cs.ubc.ca/~tmm/talks.html#cisco18

Technique-driven: Graph drawing

Daniel Archambault

David Auber (Bordeaux)

Benjamin Renoust

Guy Melançon (Bordeaux)

TreeJuxtaposer
<https://youtu.be/GdaPj8a9Q0E>

TopoLayout
 SFP
 Grouse
 GrouseFlocks
 TugGraph
<https://youtu.be/AWXAe8zvk18>

www.cs.ubc.ca/~tmm/talks.html#cisco18

Evaluation experiments: Graph drawing

Dmitry Nekrasovski

Adam Bodnar

Joanna McGrenere

Stretch and squish navigation

Jessica Dawson

Joanna McGrenere

Search set model of path tracing

www.cs.ubc.ca/~tmm/talks.html#cisco18

Technique: Dimensionality reduction

Stephen Ingram

Glimmer

DimStiller

Glimt

QSNE

www.cs.ubc.ca/~tmm/talks.html#cisco18

Evaluation experiments: Dim. reduction

Melanie Tory

Points vs landscapes for dimensionally reduced data

Michael Sedlmair

Guidance on DR & scatterplot choices

Taxonomy of cluster separation factors

www.cs.ubc.ca/~tmm/talks.html#cisco18

Evaluation in the field: Dim. reduction

DR in the Wild

Matt Brehmer

Michael Sedlmair

Melanie Tory

Stephen Ingram

www.cs.ubc.ca/~tmm/talks.html#cisco18

Curation & Presentation: Timelines

T F E



Johanna Fulda
(Sud. Zeitung)



Matt Brehmer



TimelineCurator
<https://vimeo.com/123246662>



Bongshin Lee
(Microsoft)



Benjamin Bach
(Microsoft)



Nathalie Henry-Riche
(Microsoft)

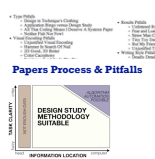


Timelines Revisited
[timelinesrevisited.github.io/](https://github.com/tmm/timelinesrevisited)

www.cs.ubc.ca/~tmm/talks.html#cisco18

Theoretical foundations

T F E



DESIGN STUDY METHODOLOGY SUITABLE

INFORMATION LOCATION

Design Study Methodology

Michael Sedlmair



Miriah Meyer



www.cs.ubc.ca/~tmm/talks.html#cisco18



Nested Model



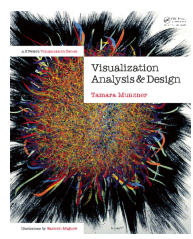
Abstract Tasks

Matt Brehmer



Theoretical foundations

T F E



Visualization Analysis & Design

Visualization Analysis & Design

www.cs.ubc.ca/~tmm/talks.html#cisco18

More information

- papers, videos, open source software, talks, courses
- <http://www.cs.ubc.ca/group/infovis>
- <http://www.cs.ubc.ca/~tmm>

www.cs.ubc.ca/~tmm/talks.html#cisco18

@tamaramunzner