DETC2013-12690

A DATA MINING TRAJECTORY CLUSTERING METHODOLOGY FOR MODELING INDOOR DESIGN SPACE UTILIZATION

Tuesday, August 6th , 2013

Yixiang Han, Conrad S. Tucker, Timothy W. Simpson, Erik Davidson.

{yzh116, ctucker4, tws8}@psu.edu
erik.davidson@buzbynetworks.com

PRESENTATION OVERVIEW

- Research Motivation
- Methodology
  - Trajectory Partitioning
  - Line Segment Clustering
- Case Study
- Conclusion and Path Forward

RESEARCH MOTIVATION

How to capture and quantify dynamics in an indoor space ?

Existing Techniques for Assessing Indoor Space Utilization

❖ Qualitative Methodologies

-Observations, interview-based and questionnaire-based surveys [1-3].

Subjective bias.

❖ Quantitative Methodologies

-Layout Optimization techniques [4-5]. E.g. Material flow optimization [5].

-Statistical Analysis [6-10]. E.g. Pyramid based methodology [6].

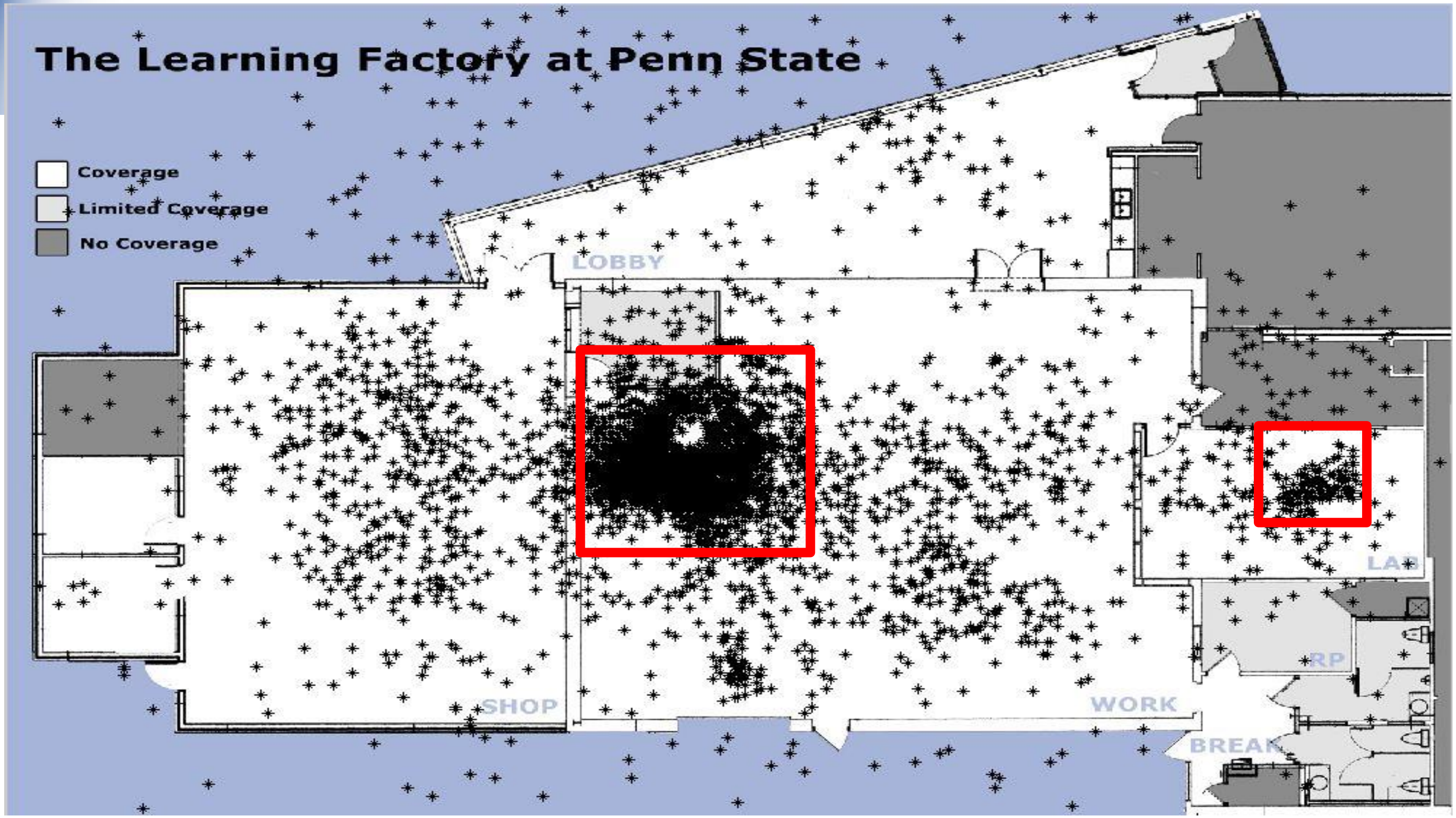
Pre-determined trajectory paths.

Research Objective

A data mining driven methodology is proposed to quantify and model common trajectory movement patterns in order to predict team dynamics and enhance indoor space design.

RESEARCH METHODOLOGY

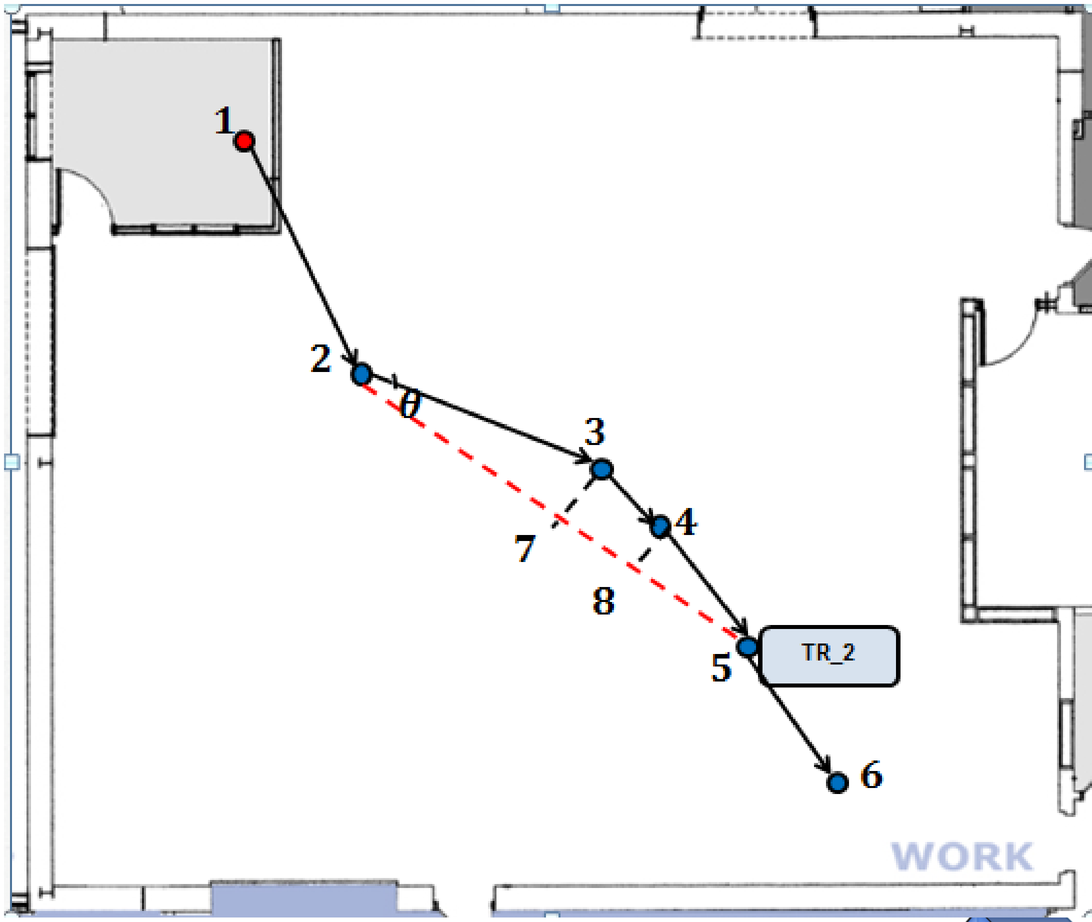
What do we get from this trajectory mapping?



Individual Trajectory Partitioning Example

Original trajectory data set: $T\_2 = \{t_1, t_2, \ldots, t_6\}$.

Characteristic point data set: P _2 $= \{p_1, p_2, p_5, p_6\}$, which is an approximation of the original trajectory.



Q: How to select characteristic point ?

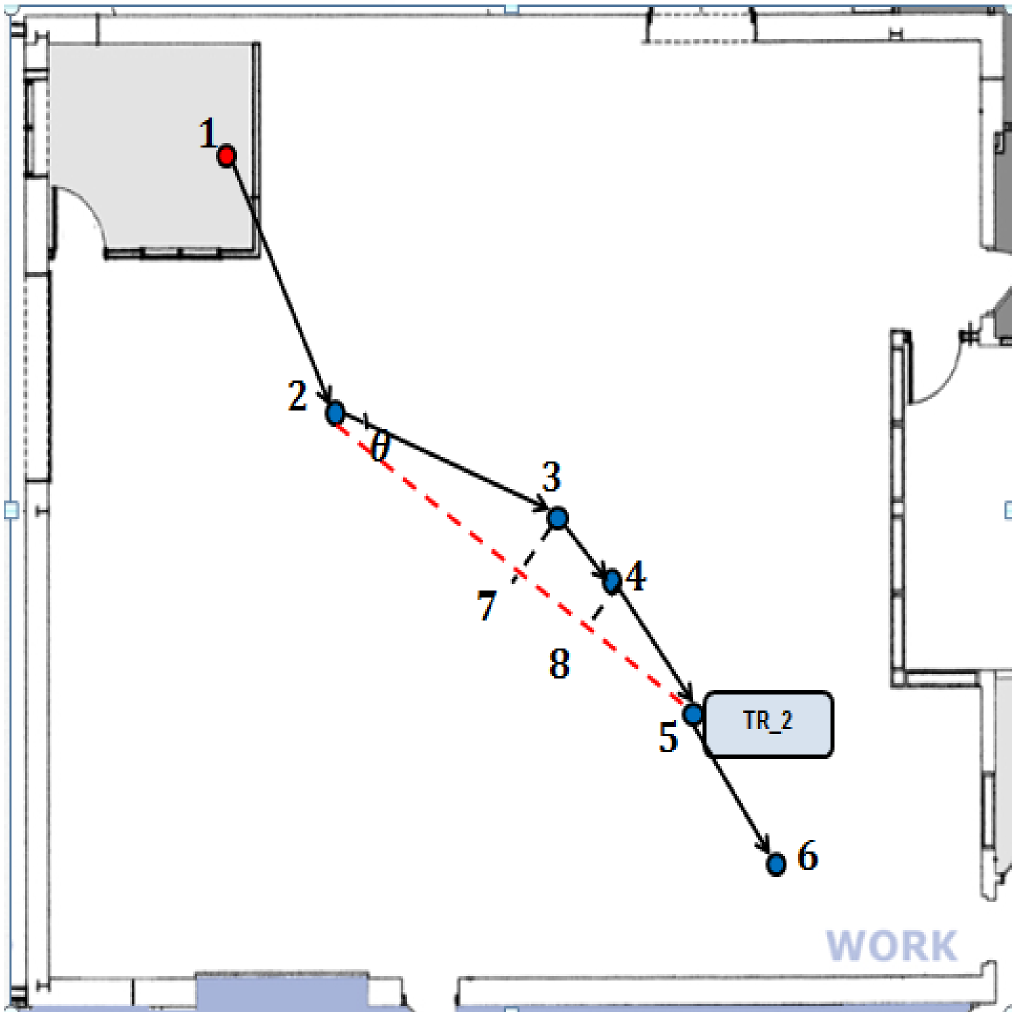
The **Minimum Description Length** (*MDL*) is applied to extract characteristic points [11].

If $MDL_{par}(t_k) > MDL_{nonpar}(t_k)$, then t_{k-1} would be a characteristic point.

e.g.

$MDL_{par}(t_5) < MDL_{nonpar}(t_5)$

$MDL_{par}(t_6) > MDL_{nonpar}(t_6)$



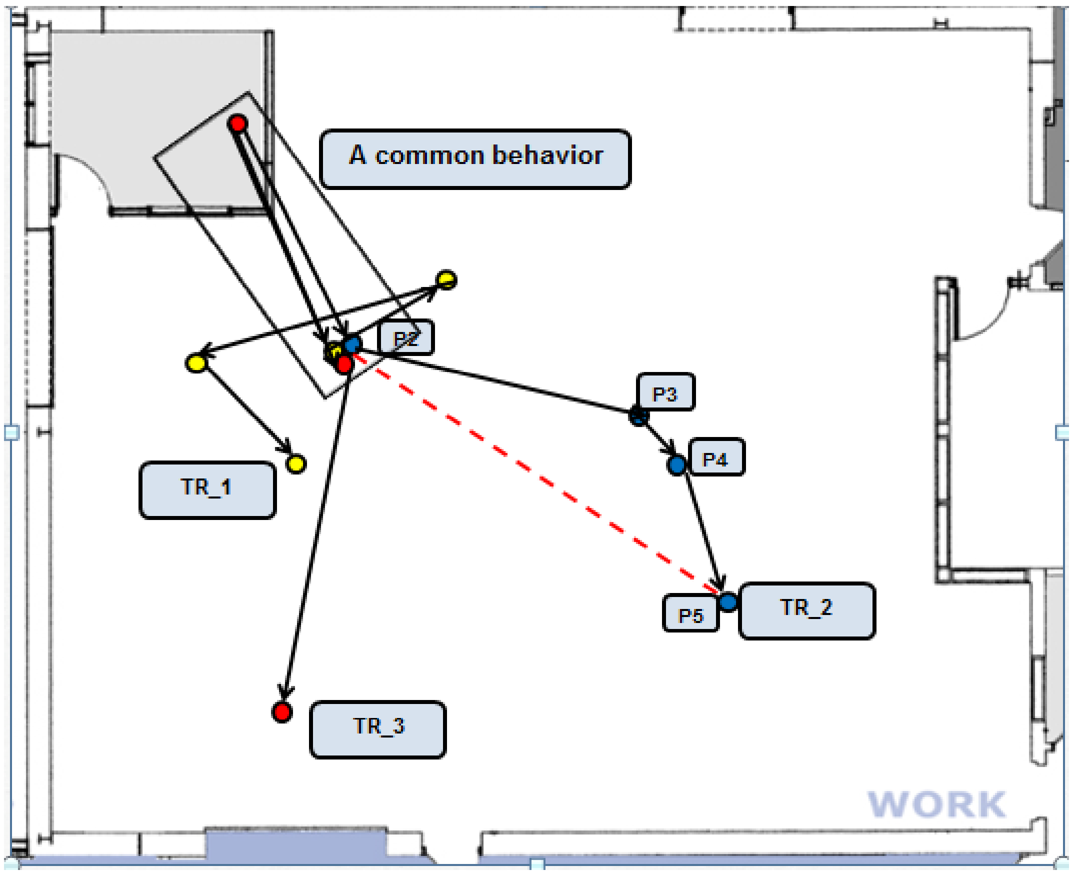
Trajectory Clustering

Given the extracted characteristic points $\mathrm{P} = \{\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_\mathrm{m}\}$from each trajectory, the clustering algorithm will group different individual movement patterns into different clusters $\mathrm{C} = \{\mathrm{c}_1, \mathrm{c}_2, \ldots, \mathrm{c}_\mathrm{n}\}$where common movement patterns are shared.

Trajectory Clustering Example

Three trajectories **TR_1, TR_2** and **TR_3** are described by characteristic points.

The line segments in the rectangular are close enough to each other, and they are considered as a cluster.

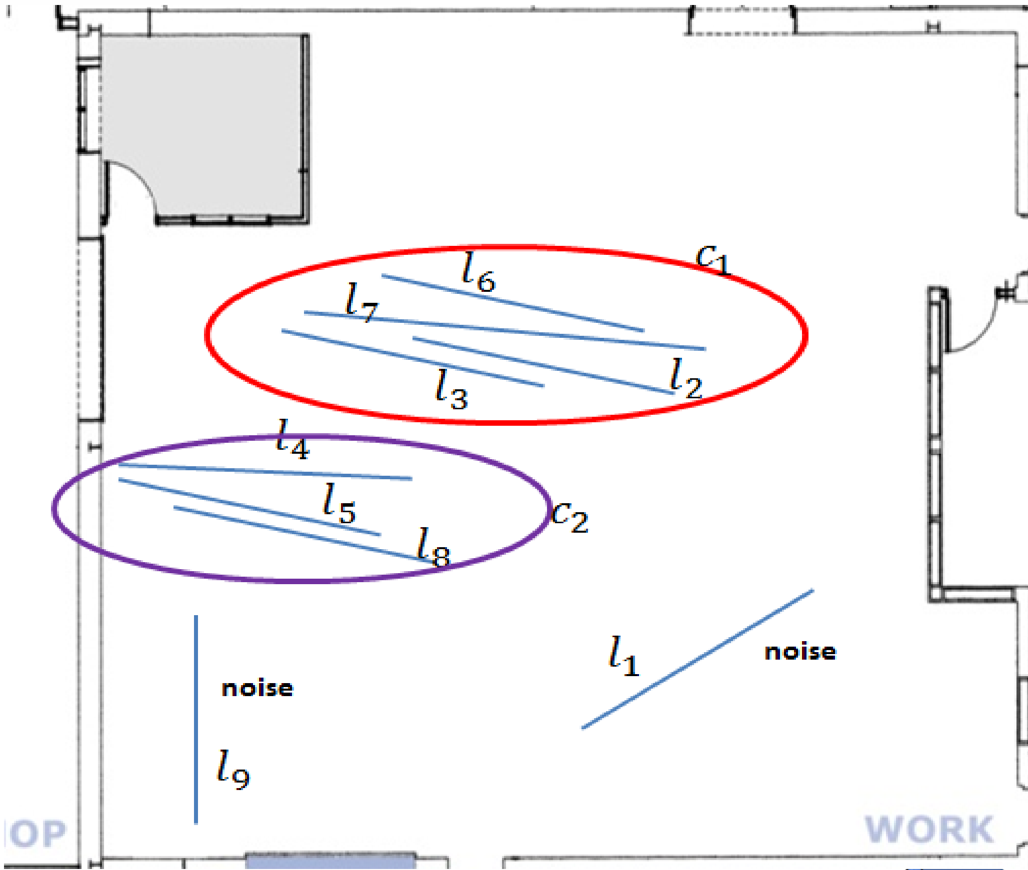


Q: How to generate cluster?

The number of neighborhood of every line segment $N_{\varepsilon}(l_i)$.

If $N_{\varepsilon}(l_i)$≥MinLns, then a density-based set is generated.

If the cardinality $(N_{\varepsilon}(l_i))$>1, then a cluster is generated.



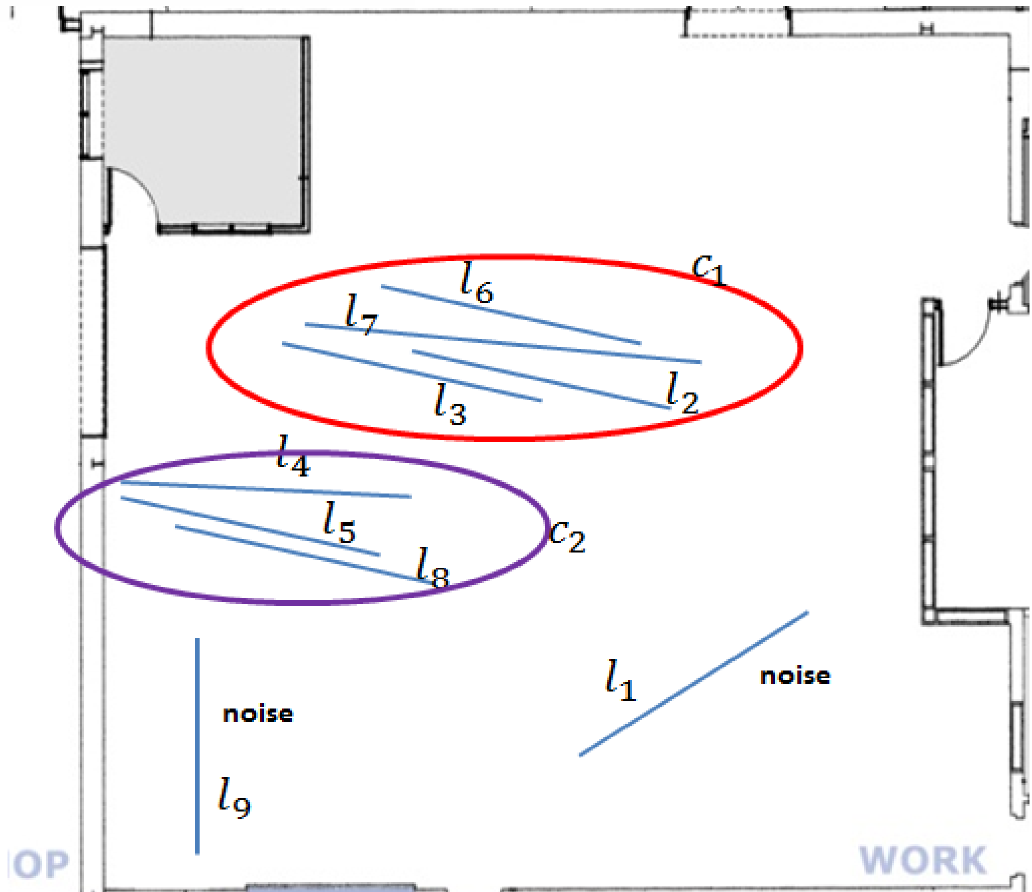
Trajectory Clustering Example

$L = \{l_1, l_2, ..., l_9\}$. Noise set $Q = \emptyset$.

Step 1:
No cluster, $L = \{l_2, ..., l_9\}$, $Q = \{l_1\}$.

Step2:
$c_1 = \{l_2, l_3, l_6, l_7\}$, $L = \{l_4, l_5, l_8, l_9\}$, $Q = \{l_1\}$.

Step 3:
$c_1 = \{l_2, l_3, l_6, l_7\}, c_2 = \{l_4, l_5, l_8\}$, L= $\emptyset$
$Q = \{l_1, l_9\}$



Learning Factory Indoor Space Utilization Case Study

Learning Factory at Penn State

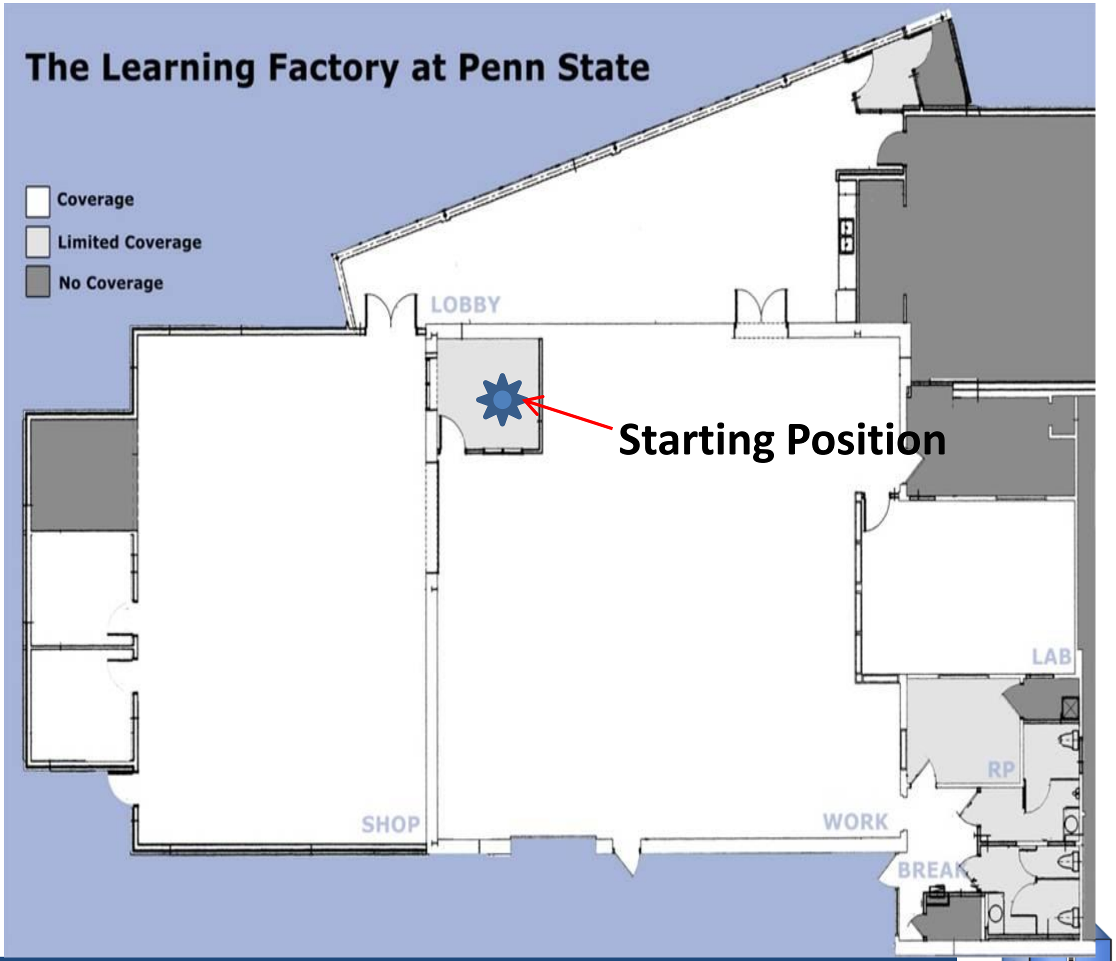
- The Learning Factory (LF) at Penn State provides student-oriented design/prototyping space for the College of Engineering at Penn State, particularly capstone design projects [12-13].

- An expansion of the LF facility in 2007 doubled its square footage; however, the program has seen explosive growth as more departments have become engaged [14].

Learning Factory Layout

Twelve tags were provided for teaching assistants (TAs).

A TA would wear one of the tags and then guide student experiments normally until the work is done and put the tag back to the container.

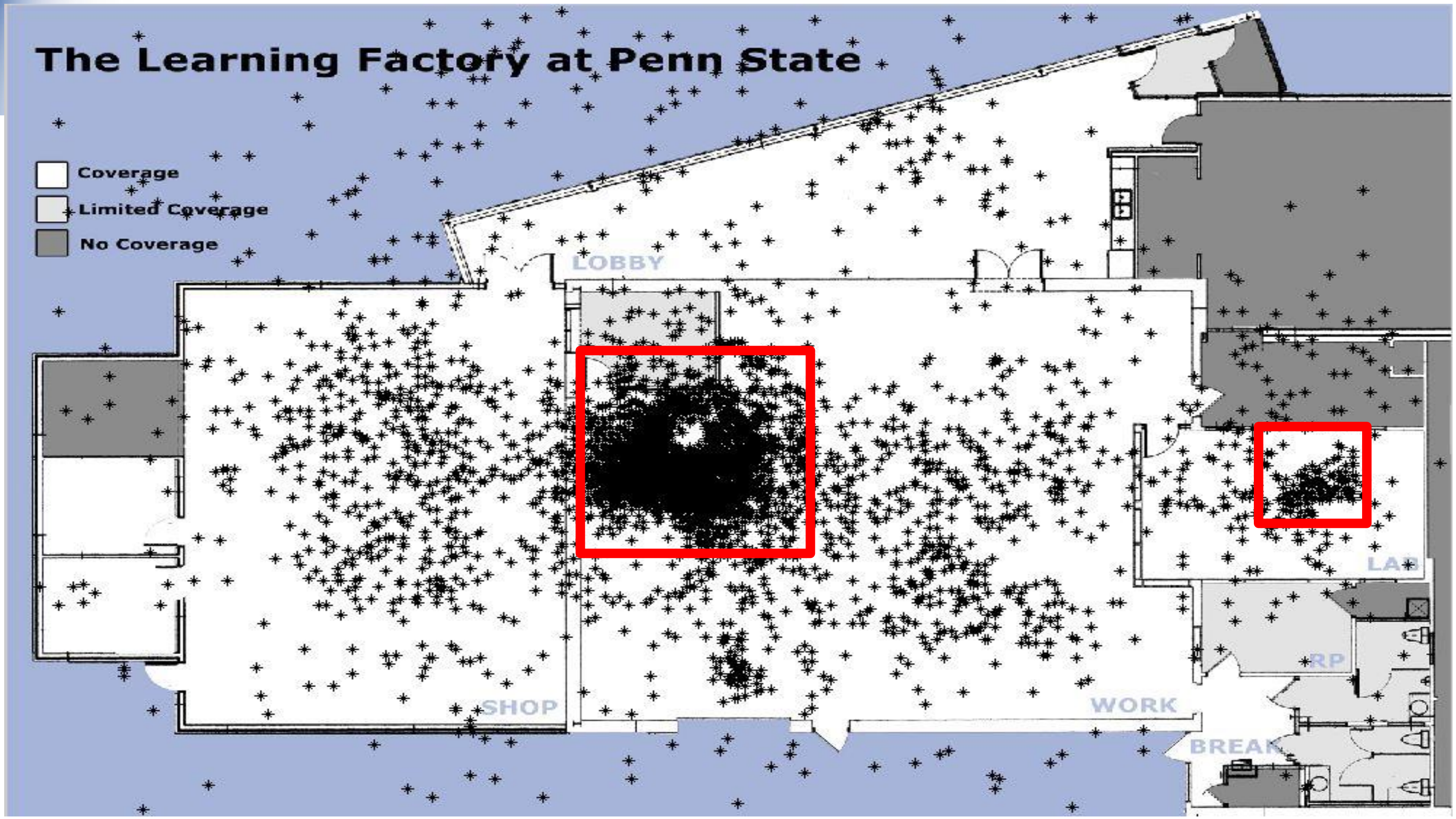


Clustering Result

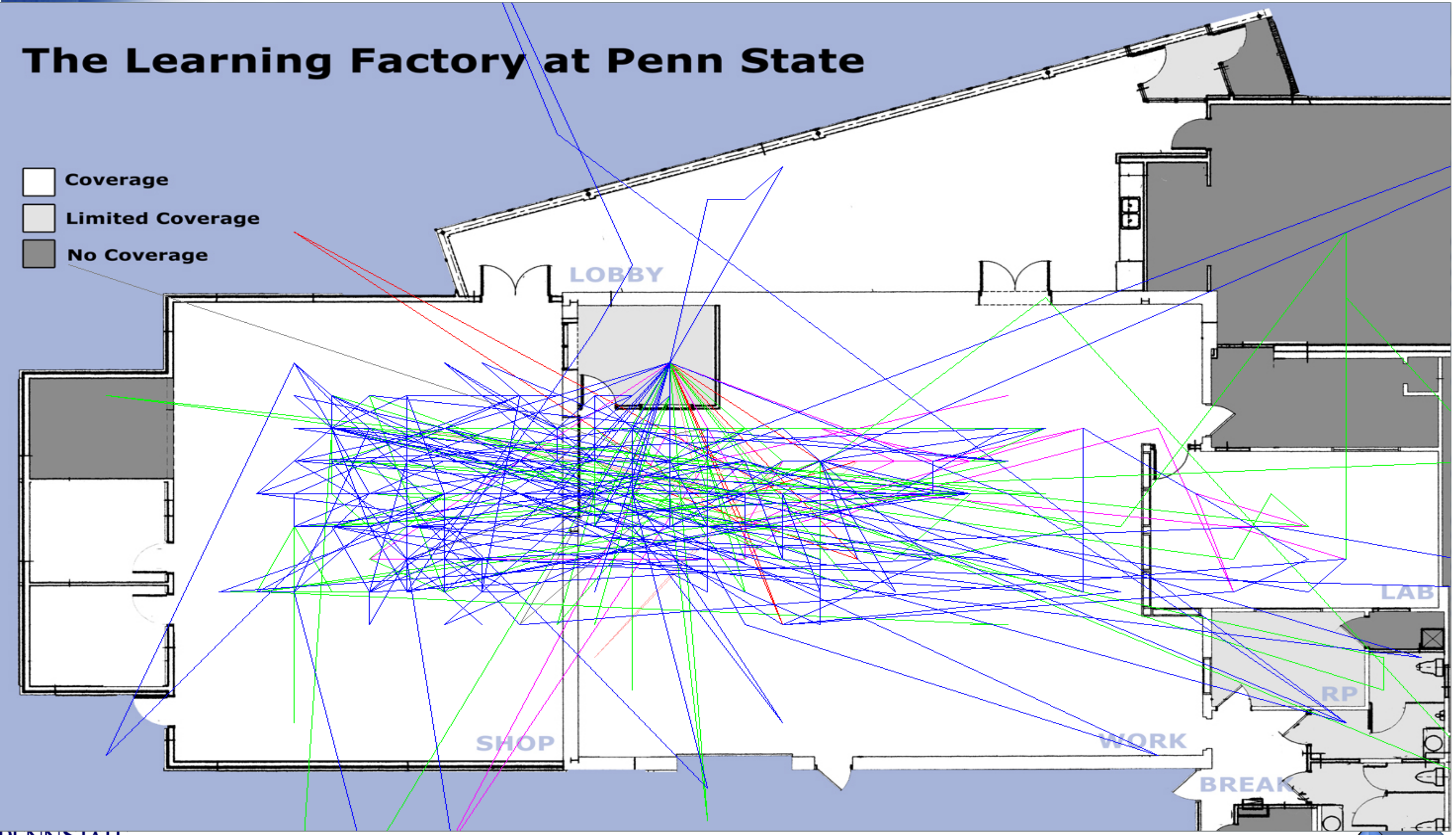
TABLE 2. Statistical clustering result

Cluster No.	Total number of line segments	Cluster cardinality
C1	58	20
C2	41	18
C3	8	3
C4	42	15
C5	15	8
C6	59	14
C7	48	14
C8	224	46
C9	322	44

What do we get from this trajectory mapping?



Visualization



Results:

1. Nine types of common movement patterns are generated.

2. Cluster 8 and Cluster 9 can explain the most significant movement patterns as large number of individuals are included. At the same time, we can see the “back-and-forth” pattern represented.

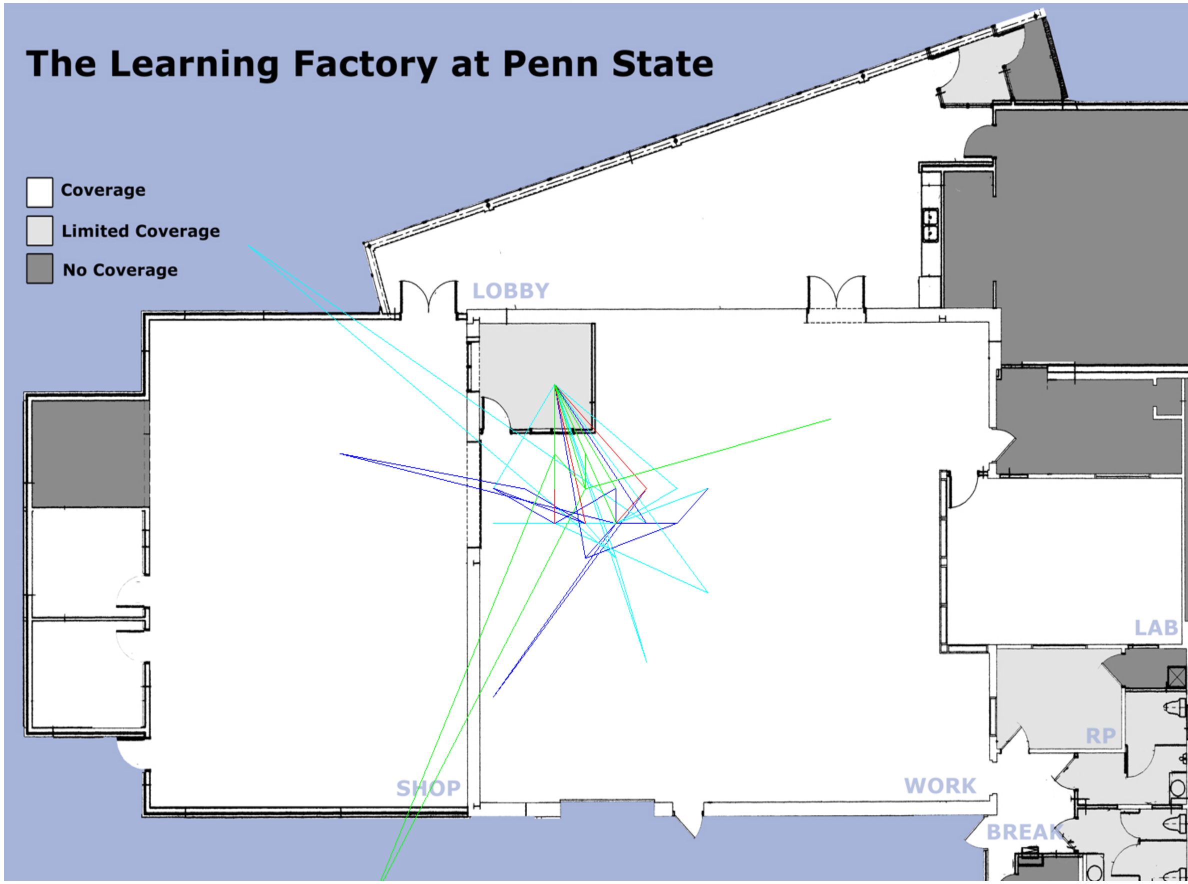
3. Two middle spaces are most utilized regions. In addition, they are always utilized simultaneously in the Learning Factory.

Movement Pattern Evolution Detection

The objective: detect any change of indoor space utilization patterns in the Learning Factory.

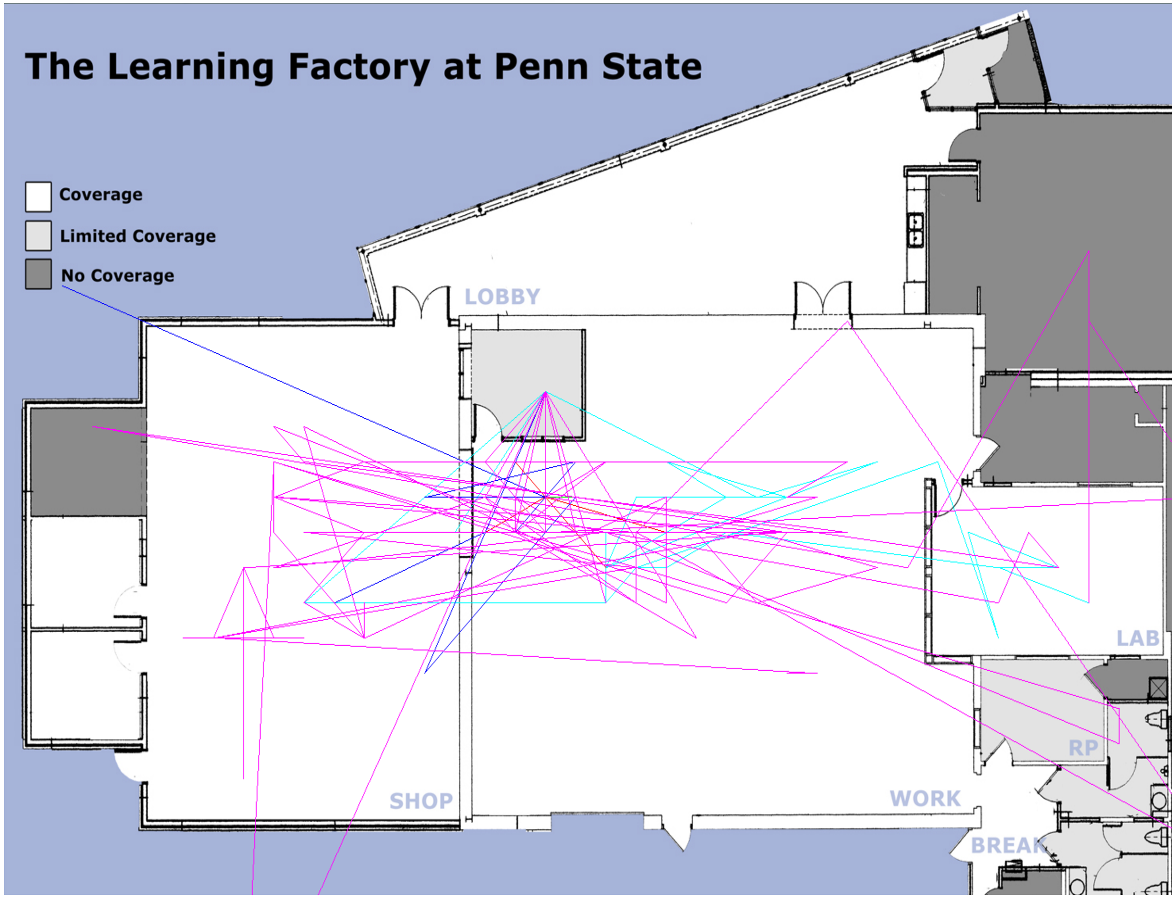
Visualization of Period I (01/20/2012 to 02/21/2012)

- Four clusters.
- Part of the middle spaces are utilized.



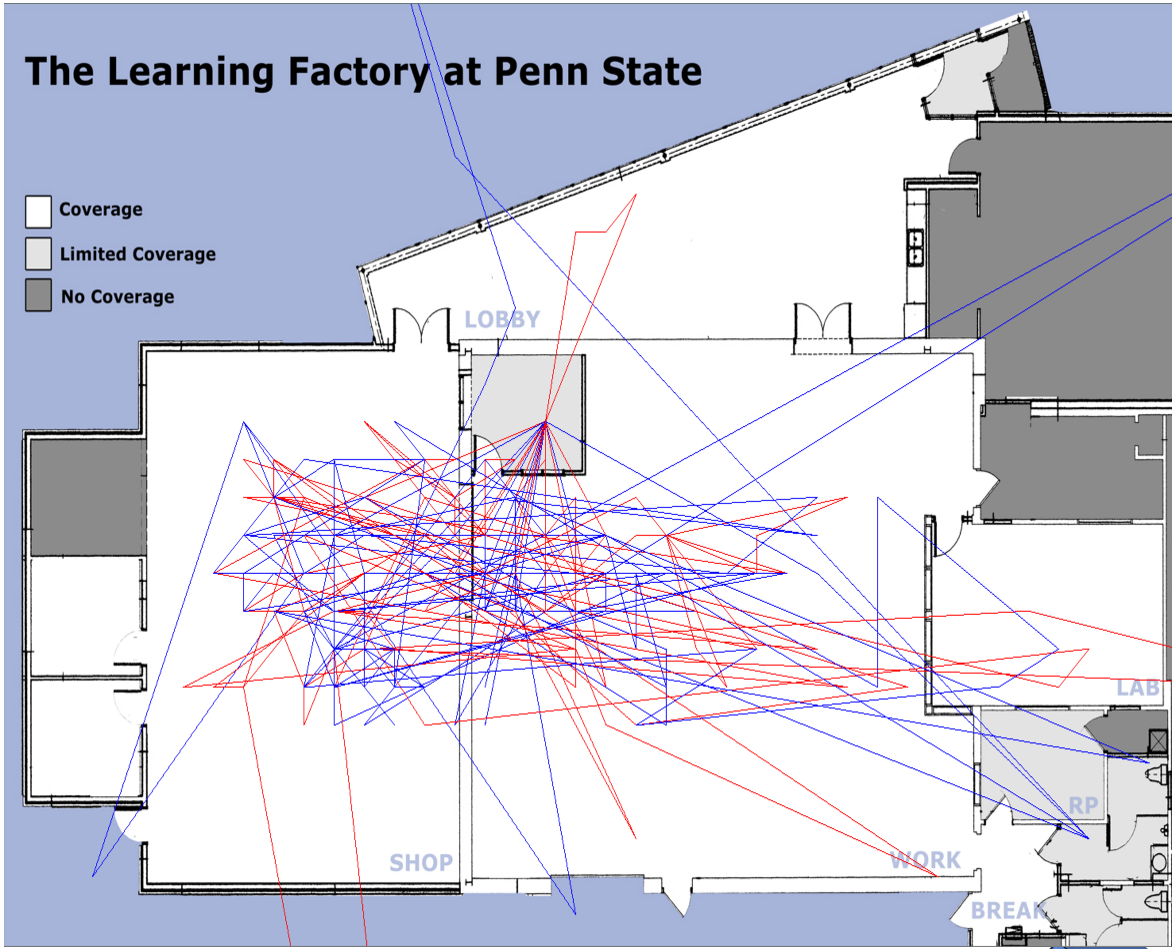
Visualization of Period II (02/22/2012 to 03/22/2012)

- Four clusters.
- More spaces are utilized including Lab Space and Toilet.



Visualization of Period III (03/23/2012 to 04/23/2012)

- Only two clusters which means the trajectory patterns are tend to be more similar comparing to the first two periods.



Results:

4. Utilized spaces are increasing from the Period I to the Period 3.

5. The similarities among multiple clusters are increasing from the Period I to the Period 3 since the number of clusters are decreasing from Period I to Period III.

Conclusion:

In this paper, we propose a data mining driven methodology which is able to model and predict common trajectory movement patterns in order to understand team dynamics and navigate indoor space design.

Future Work

Future Work

1. Include indoor facility layout optimization to enhance team dynamics and overall project quality ;

2.Explore other potential indoor space design applications such as emergency room in hospital, etc.

Acknowledgement

Contributors:

- Yixiang Han, Conrad S. Tucker, Yixiang Han, Timothy W. Simpson, Erik Davidson.

References:

[1] Tatsuya Nomura, Takugo Tasaki, Takayuki Kanda, Masahiro Shiomi, Hiroshi Ishiguro, and Norihiro Hagita. Questionnaire-based research on opinions of visitors for communication robots at an exhibition in japan. In Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction, INTERACT'05, pages 685–698, Berlin, Heidelberg, 2005. Springer-Verlag.

[2] T. Shibata, K. Wada, and K. Tanie. Tabulation and analysis of questionnaire results of subjective evaluation of seal robot at science museum in london. In Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on, pages 23–28. IEEE, 2002.

[3] F. Girardin, F. Dal Fiore, C. Ratti, and J. Blat. Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. Journal of Location Based Services, 2(1):41–56, 2008.

[4] B. Benson and BL Foote. Doorfast: A constructive procedure to optimally layout a facility including aisles and door locations based on an aisle flow distance metric. International Journal of Production Research, 35(7):1825–1842, 1997.

[5] SP Singh and R.R.K. Sharma. A review of different approaches to the facility layout problems. The International Journal of Advanced Manufacturing Technology, 30(5):425–433, 2006.

References Continued:

[6] Y. Yoshimura, F. Girardin, J.P. Carrascal, C. Ratti, J. Blat, M. Fuchs, F. Ricci, L. Cantoni, et al. New tools for studying visitor behaviors in museums: a case study at the louvre. In Information and Communication Technologies in Tourism 2012. Proceedings of the International conference in Helsingborg (ENTER 2012): Springer Wien New York. This entry was posted on February, volume 6. in: ENTER12 19th International Conference on Information Technology and Travel & Tourism ENTER 2012, 2012.

[7] G. Andrienko, N. Andrienko, and S. Wrobel. Visual analytics tools for analysis of movement data. ACM SIGKDD Explorations Newsletter, 9(2):38–46, 2007.

[8] H. Cao, N. Mamoulis, and D.W. Cheung. Mining frequent spatial-temporal sequential patterns. In Data Mining, Fifth IEEE International Conference on. IEEE, 2005.

[9] S. Rinzivillo, D. Pedreschi, M. Nanni, F. Giannotti, N. Andrienko, and G. Andrienko. Visually driven analysis of movement data by progressive clustering. Information Visualization, 7(3-4):225–239, 2008.

[10] S. Kisilevich, F. Mansmann, M. Nanni, and S. Rinzivillo. Spatial-temporal clustering: a survey. In Proceedings of the 2007 ACM SIGMOD international conference on Management of data. Technical Report, ISTI-CNR, Italy, Submitted to Data Mining and Knowledge Discovery Handbook, 2007.

[11] J.G. Lee, J. Han, and K.Y. Whang. Trajectory clustering: a partition-and-group framework. In Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pages 593–604. ACM, 2007.

[12] J.S. Lamancusa. The reincarnation of the engineering shop. In ASME Design Engineering Technical ConferencesDesign Education Conference. ASME, 2006.

[13] J.S. Lamancusa and T.W. Simpson. The learning factory–10 years of impact at penn state. In International Conference on Engineering Education, pages 16–21, 2004.

References Continued:

[14] T.W. Simpson and E. Kisenwether. Driving entrepreneurial innovation through the learning factory: The power of interdisciplinary capstone design projects. In ASME Design Engineering Technical Conferences-Design Education Conference. ASME, 2013.

Thank you

Q &A