Online Urbanism: Interest-based Subcultures as Drivers of Informal Learning in an Online Community

Ben U Gelman, Chris Beckley, Aditya Johri, Carlotta Domeniconi, and Seungwon Yang*

*George Mason University, *Louisiana State University*

Acknowledgements:
We appreciate the Lifelong Kindergarten group at MIT for publicly sharing the Scratch datasets. This work is partly based upon research supported by U.S. National Science Foundation (NSF) Award#DUE-1444277 & EEC-1408674

Purpose

- Understand the growth of online communities of creators over time
- Observe the effects of population growth on user interests and network structure
- Place our research in the context of real life urbanism as well as online community design that supports an evolving population

Overview

- Review terms – informal learning and online communities of creators
- What is Scratch? Data and community structure
- Clustering method and validation
- Temporal analysis of clusters

Informal Learning

- Contrasts a classroom setting
- “Predominately unstructured, experiential, and non-institutional”
  - Victoria Marsick*
- Lack of standardized examination or measurement

*Marsick, V. J., & Volpe, M. (1999). The nature of and need for informal learning. In V. J. Marsick & M. Volpe (Eds.), *Advances in developing human resources: Informal learning on the job* (pp. 1–9). Baton Rouge, LA: AHRD.

How I Learned to Type–Runescape 2004-2006

Online Communities of Creators

- A specific type of informal learning community

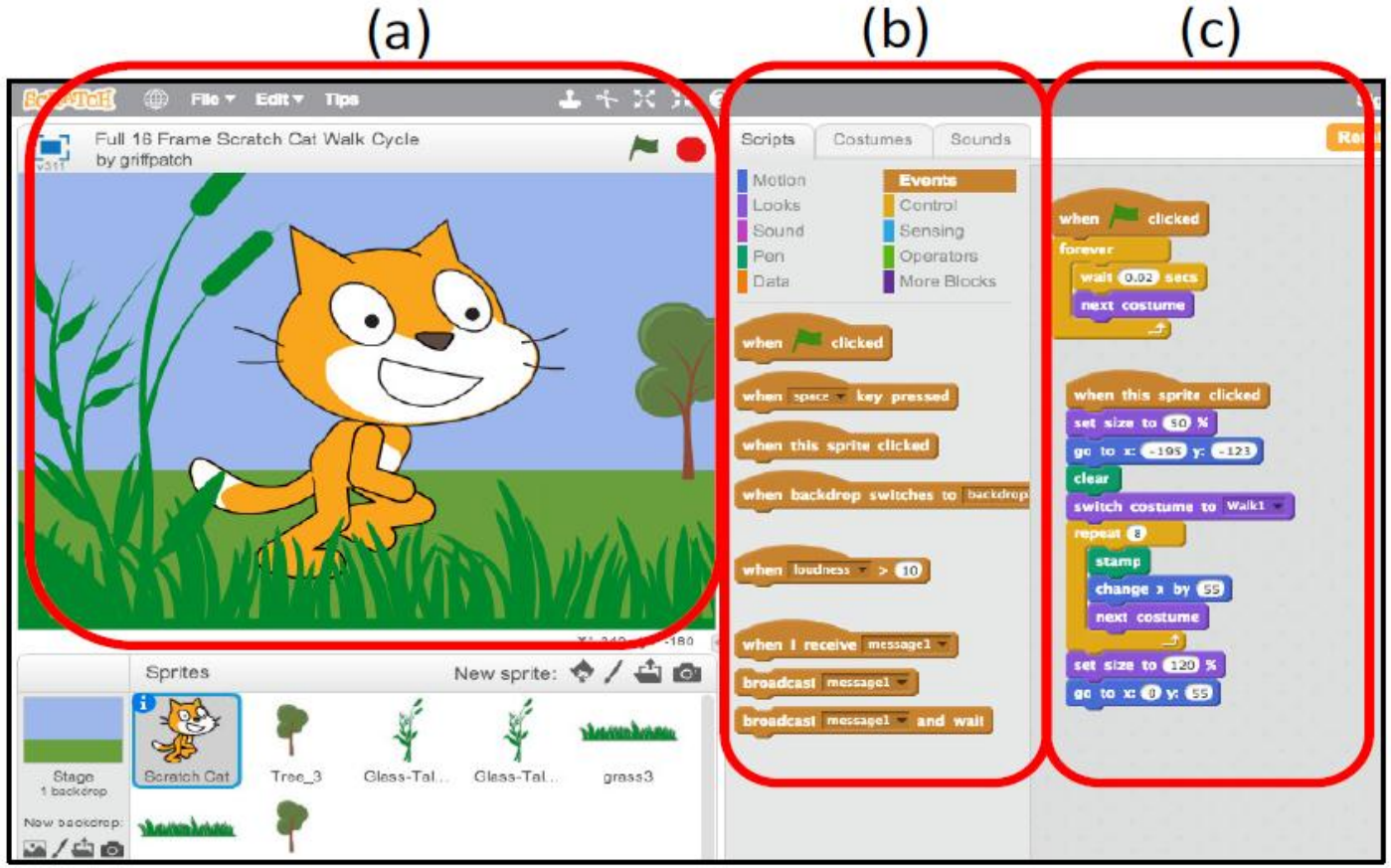
- Social network sites where the core activity of members is sharing personal and original creations

Scratch

- An online community of creators – our focus in this study
- Visual programming interface
- Users are able to:
  - Post projects
  - View and edit source code of other projects – “remixing”
  - Follow other users
  - Chat, comment, and post on the forums

Scratch Project Interface

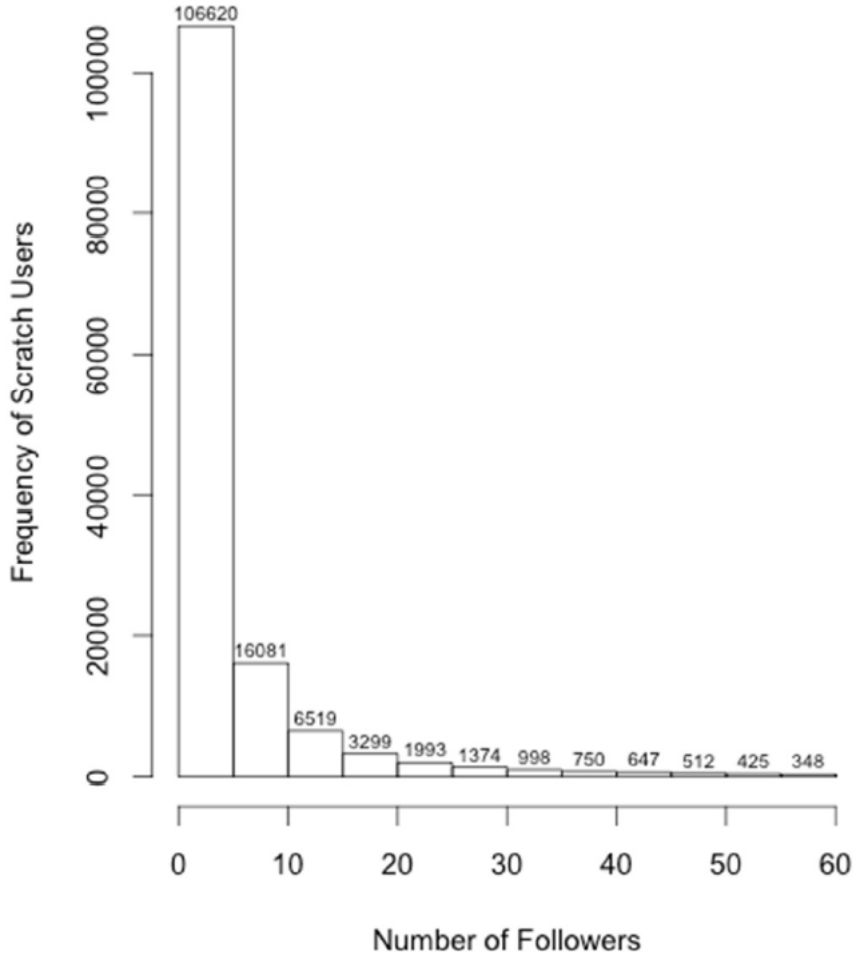
- (a) – Project window
- (b) – Block selection
- (c) - Script

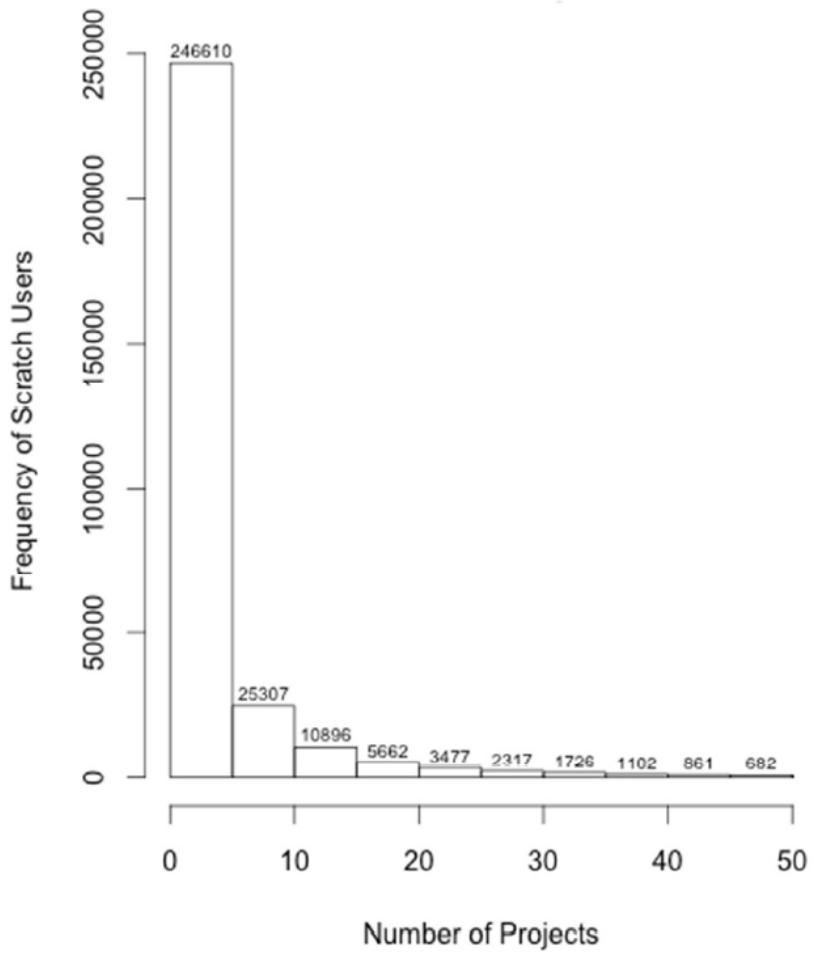


Scratch Data

- March 2007 – April 2012
- 1,056,950 registered users
- 1,928,699 projects
- 170 Programming blocks

Scratch Data: Long-tail Phenomenon

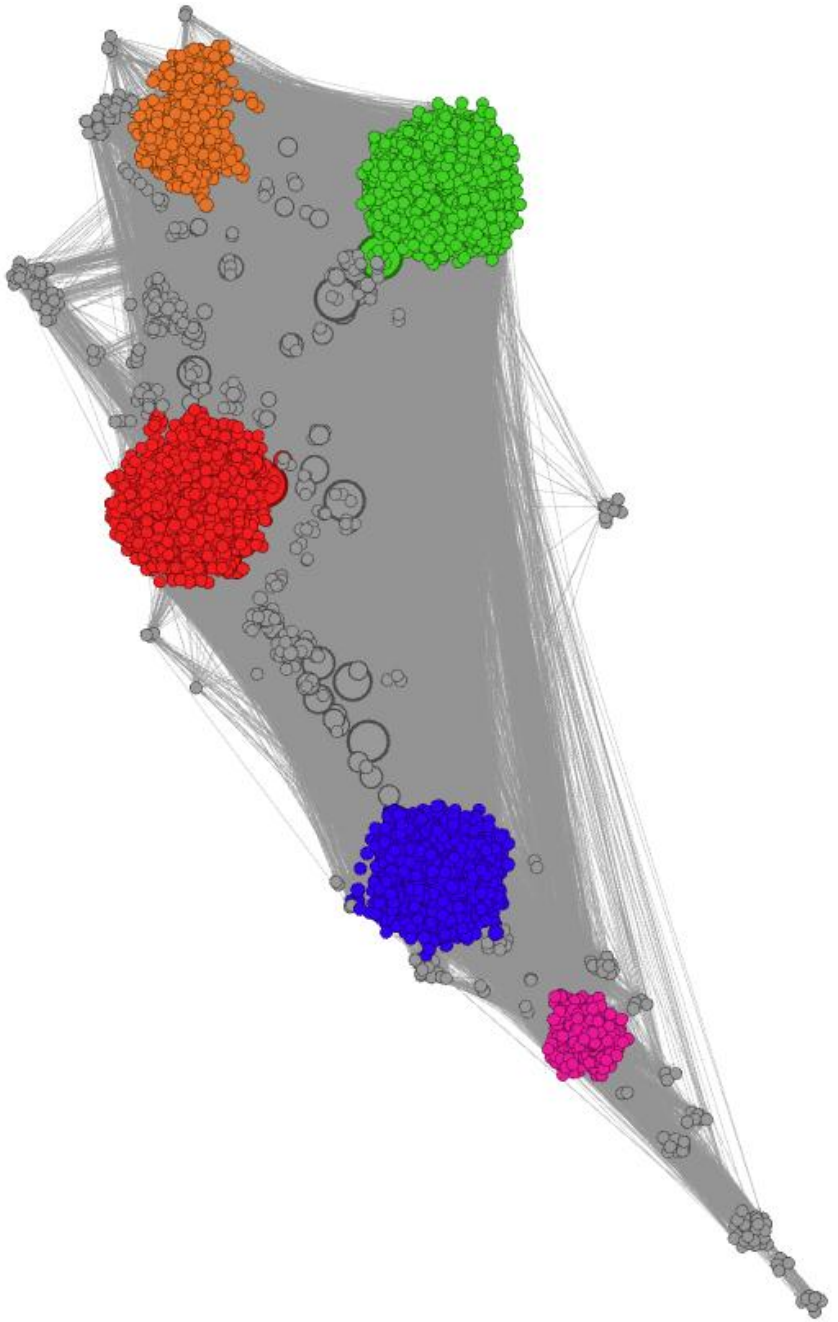




Community Clustering

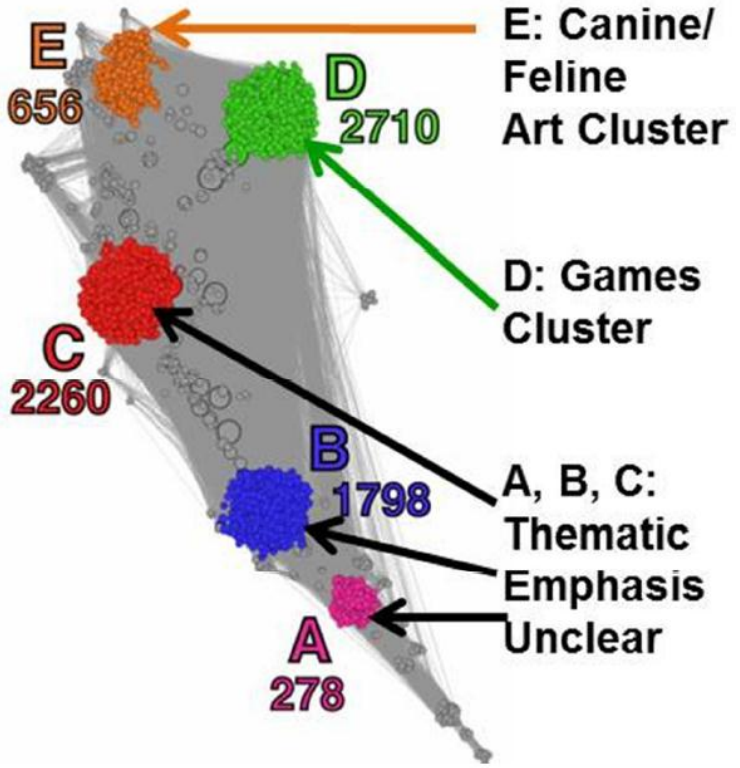
- Users with 25 followers
- Gephi
  - Graphing and visualization tool
- OpenOrd
  - Multi-level
  - Average-link clustering

Clustering Results



Validating Cluster Themes

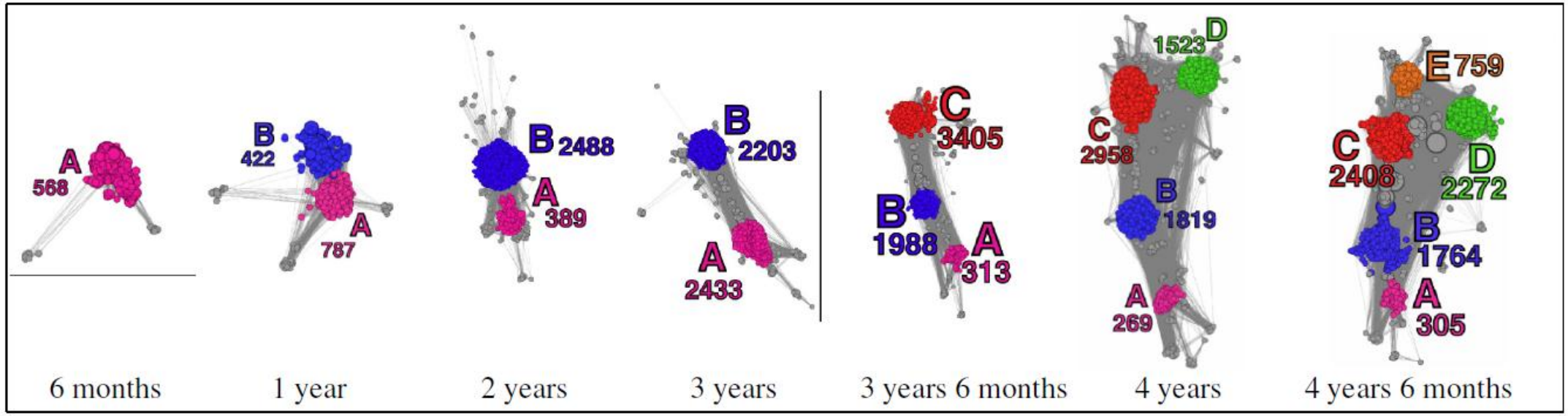
- Text mining approach
- TFIDF
  - Term Frequency Inverse Document Frequency



Cluster	Top Words (TFIDF Values as percentage of Total Terms)
A	comput (5.3); time (4.7); make (4.4); space (4.4); random (4.0); new (3.8); stupid (3.6); first (3.2)
B	waffle (19.3); mario (17.6); game (12.3); super (9.5); tag (9.3); fun (8.5); pokemon (8.0)
C	funni (16.2); waz (15.0); wuz (15.0); tag (11.8); pokemon (10.4); good (8.6); add (8.6); sonic (8.5)
D	mario (22); epic (20.7); tag (16.3); press (12.2); sonic (9.8); super (8.4); add (6.2); scratch (4.2)
E	warrior (11.5); wolf (9.6); yay (8.8); cat (7.9); contest (7.7); pleas (7.0); art (5.7); pie (5.7)

Temporal Emergence of Subgroups

- Sliding window approach
- Each window shows users from the start of the data set up to some point in time after the start
- Each window clustered by OpenOrd



Scratchage and Scratchend

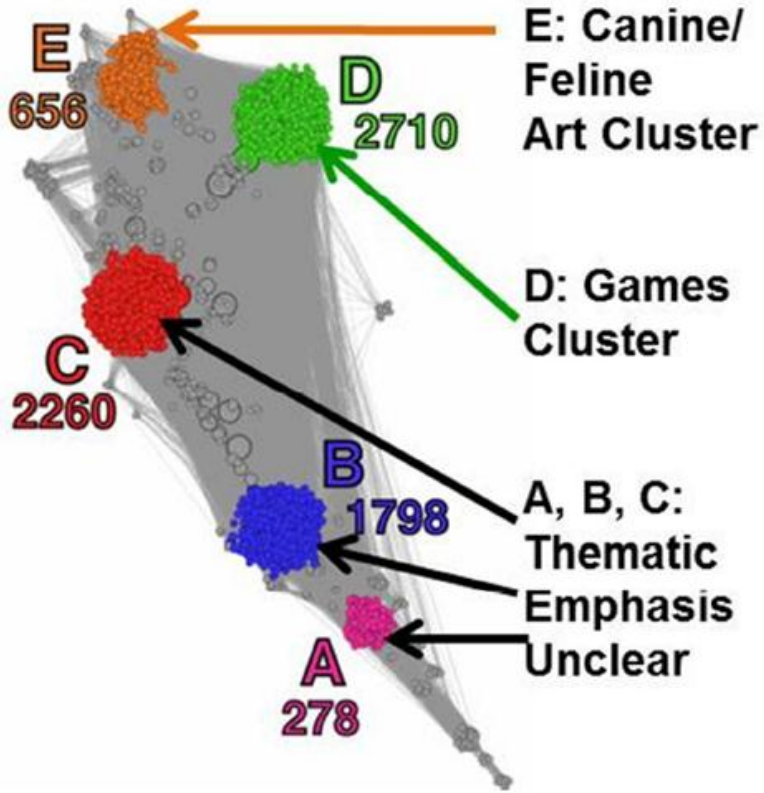
- Scratchage – Age of a user in years starting from March 5, 2007
  - A user who created their account on March 5, 2007 will have a Scratchage of 1 on March 5, 2008.

- Scratchend – Time in years between a user's final activity and the end of the data set on April 1, 2012
  - A "final activity" is either following another user or posting a project.
  - A user whose final activity was on April 1, 2011 will have a Scratchend of 1.

Cluster Scratchages and Scratchends

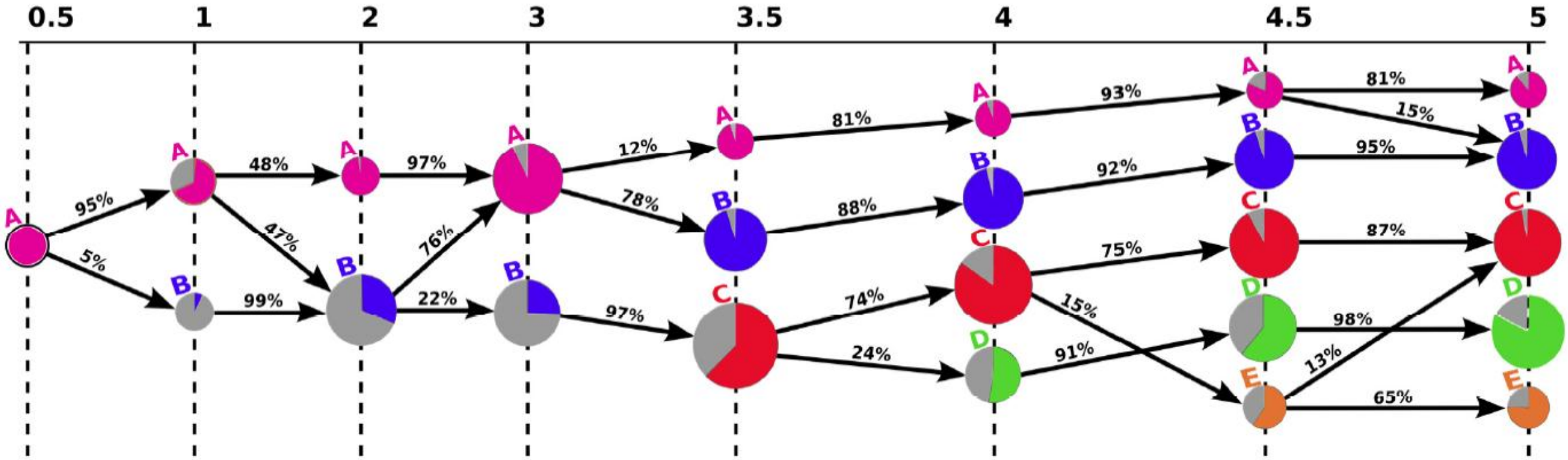
- Time was not a factor in the clustering
- Temporal trends emerged anyways

Cluster	SA: Avg.	SA: SD	SE: Avg.	SE: SD
A	4.85	0.25	3.80	1.28
B	4.11	0.46	2.43	1.25
C	2.79	0.68	1.19	0.86
D	1.69	0.89	0.37	0.46
E	1.25	0.72	0.23	0.30



Migration Patterns

- Size of circle is proportional to size of cluster
- Grey sections of each circle are new users



Limitations and Future Work

- Lack of demographic data
- Cumulative Analysis
- Future work: individual user development

Discussion

- Fischer's* Subcultural Theory of Urbanism
- Views subcultures as a "set of interconnected social networks"
- Argues that population growth provides opportunities for similar people to group

Fischer, C. S. (1975). Toward a subcultural theory of urbanism. *American journal of Sociology*, 1319-1341.
Fischer, C. S. (1995). The subcultural theory of urbanism: A twentieth-year assessment. *American Journal of Sociology*, 543-577.

Discussion

- Support newcomers by targeting their interests, help them find place in the community
- Normally done by curating interesting projects, targeting user groups or experience levels
- Interest-based clusters allow for strategic project curation

Online Urbanism: Interest-based Subcultures as Drivers of Informal Learning in an Online Community

Ben U Gelman, Chris Beckley, Aditya Johri, Carlotta Domeniconi, and Seungwon Yang*

*George Mason University, *Louisiana State University*

Acknowledgements:
We appreciate the Lifelong Kindergarten group at MIT for publicly sharing the Scratch datasets. This work is partly based upon research supported by U.S. National Science Foundation (NSF) Award#DUE-1444277 & EEC-1408674