

VisPaD: Tool for Visualization and Pattern Discovery for Fighting Human Trafficking



Pratheeksha Nair^{†,‡}, Yifei Li^{†,‡}, Catalina Vajiac^{*}, Andreas Olligschlaeger^{†,‡}, Meng-Chieh Lee^{*}, Namyong Park^{*}, Duen Horng Chau[¶], Christos Faloutsos^{*}, Reihaneh Rabbany^{†,‡}

[†]McGill University; [‡]MILA; ^{*}Carnegie Mellon University; [¶]Georgia Institute of Technology; [§]Marinus Analytics

Background (Why?)

- × Human Trafficking (HT) affects over 4.8 million people world-wide
- × Majority of victims are advertised online
- × One trafficker writes ads for multiple victims
- × HT detection algorithms look for clusters of similarly phrased ads

Problem (What?)

Given the large number of clusters, how can we help expedite the process of identifying interesting ones?

How can we visualize these clusters to draw useful insights from them?

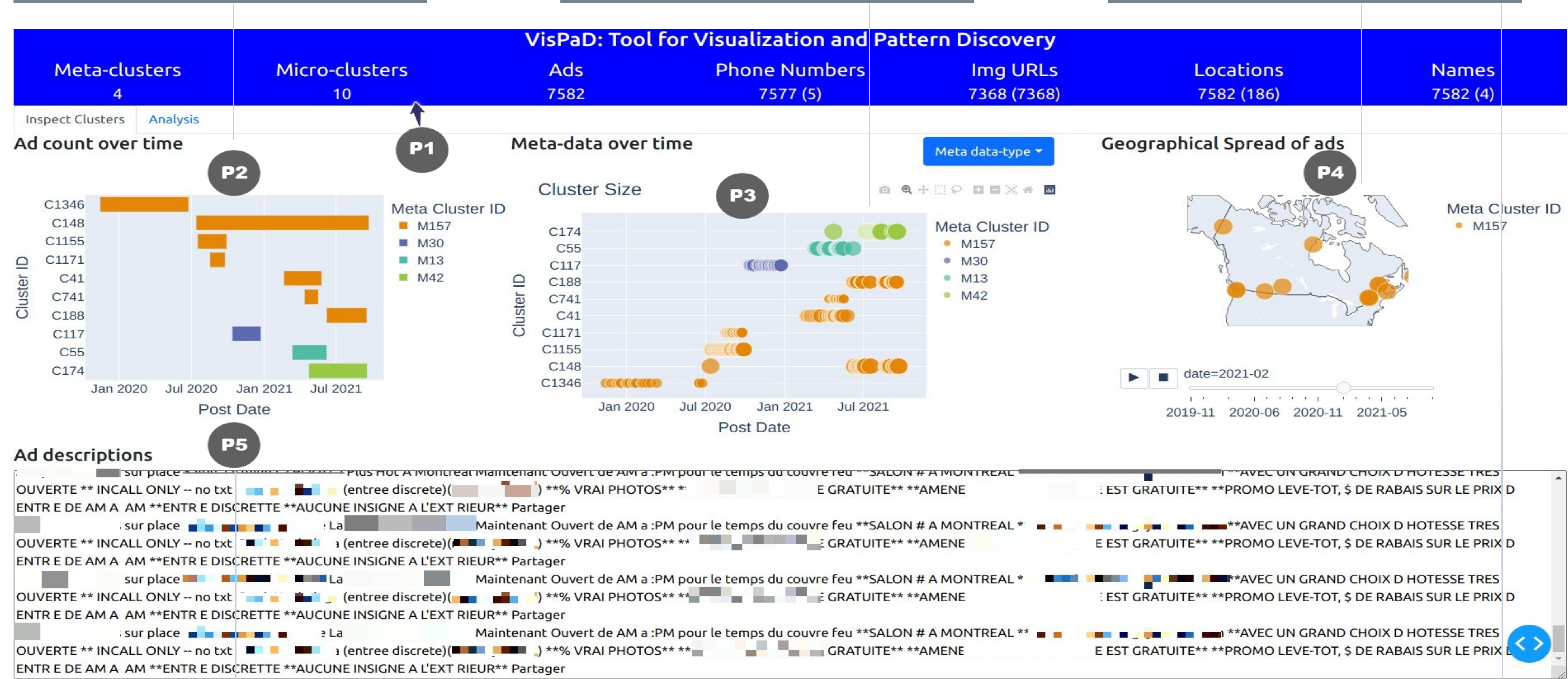
Proposed tool (How?)

- ✓ **Characterize the clusters** and represent them as embeddings
- ✓ **Interactively visualize** the cluster embeddings
- ✓ Identify **interesting patterns**
- ✓ Select the interesting clusters and **inspect their spatio-temporal information**

P2: View cluster activity over time. Which ones post for longer duration?

P3: Inspect meta-data behavior. Is there an increase in number of images?

P4: Identify spatio-temporal hotspots. Are the ads posted all over the country?

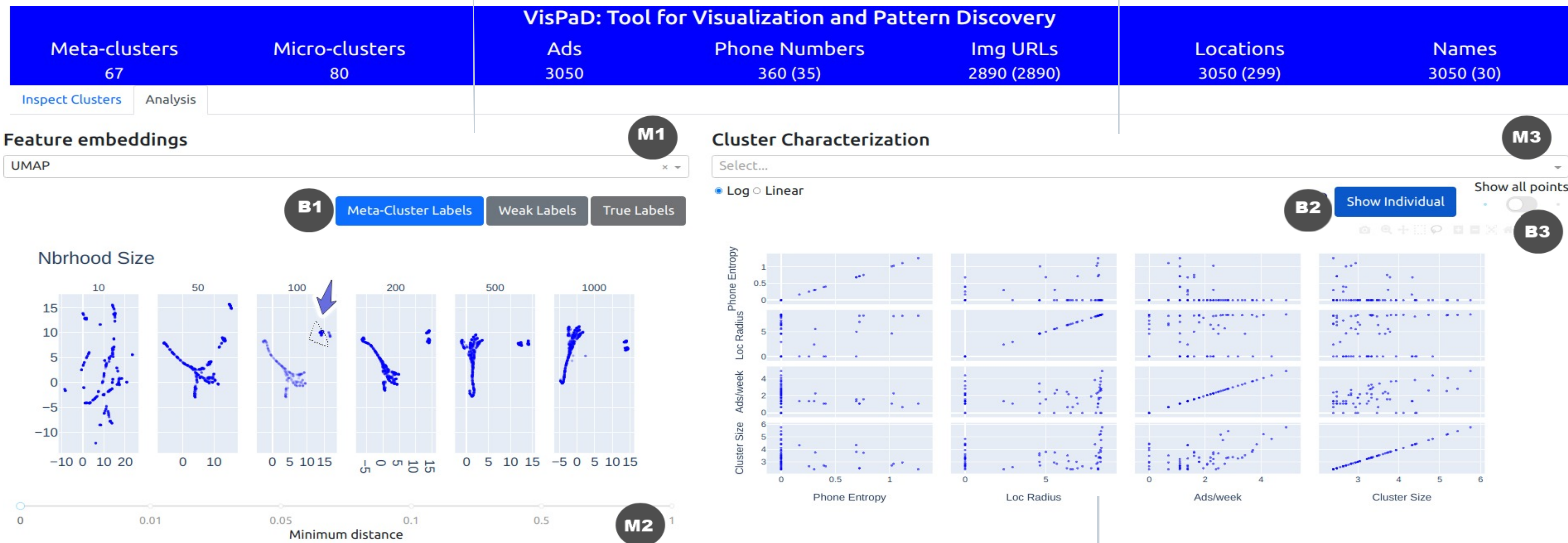


P5: Look at patterns in the ad text. Are the ads posted all over the country?

P1: Summarized info of selected clusters and associated meta-data

Choose method of visualizing cluster embeddings

Select particular features to inspect



Select interesting pattern of clusters for inspection

Which features separate the selected clusters?