Using Similarity Measures to Detect Organizations in Online Escort Advertisements

Carl Edwards, Anthony Wertz, and Artur Dubrawski

Motivation

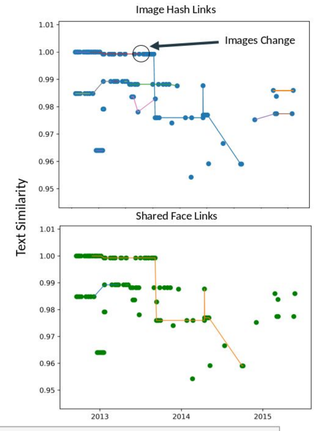
- Human trafficking is still a pervasive and global problem today
- 40 million people globally are involved in modern slavery [1]
- This activity has moved online in the form of social networking and online classifieds [2]
- The dataset consists of roughly 40 million advertisements from the escorts section of backpage.com from 2012 to 2017
- The usage of similarity measures can incorporate multiple modalities to allow for detection and monitoring of organizations

Similarity Measures

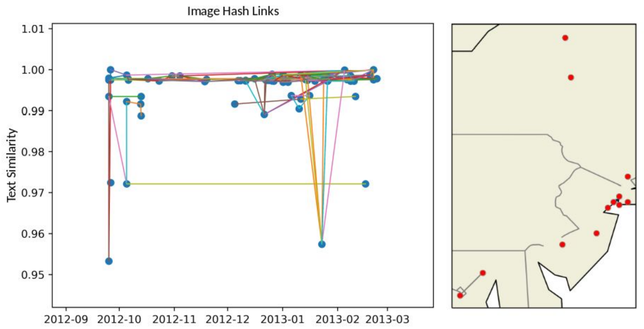
- **Text Similarity**:
  - Train fastText on corpus of advertisements
  - Generate paragraph embeddings of advertisements as average of constituent word vectors
  - Use cosine similarity between paragraph embeddings
- **Common-Feature Similarity**: Do two ads share some feature in common
  a. **Phone Number Similarity**
  b. **Image Hash Similarity**
  c. **Name Similarity**
- **Face Similarity**
  - Face recognition pipeline implemented using DLIB:
    - CNN face detection
    - 5-point landmarking and face localization
    - Creation of face 128D face embedding using ResNet with 29 convolutional layers

Results

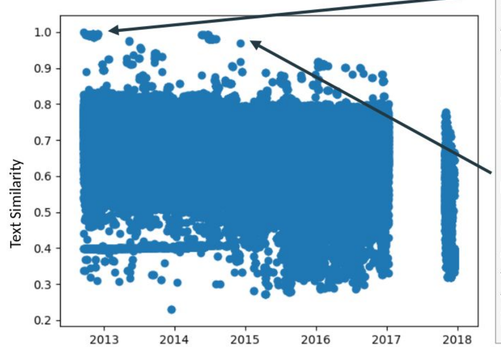
- Cosine similarity between unsupervised paragraph embeddings can be used to extract related ads in noisy text data.
- Additional similarity measures can be used to link organizations together amidst noise
- Multiple modalities can be used to improve the connection between ads
- Organizations can be monitored within single cities and on a country-wide level



Location and Image



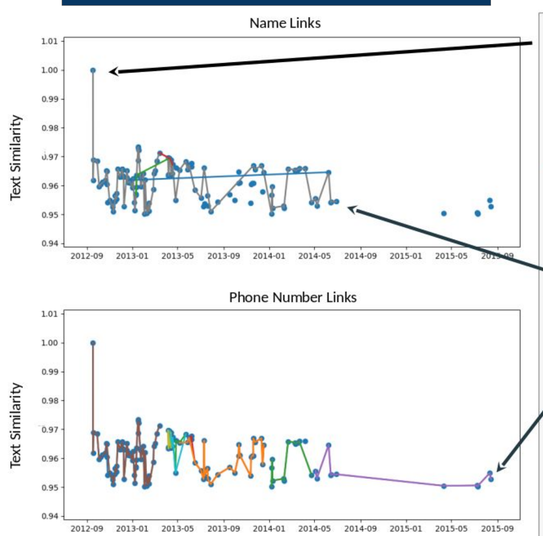
Text Similarity



September 2012:
???:*??? Hi, I'm MEGAN ???:*??? ???:*??? For those guys who expect the BEST, stop searching My services are of TOP quality and I can guarantee complete satisfaction. ???:*??? ???:*???Take a break from your stressful lifestyle and enjoy a fun-filled relaxing time with me.... ???:*??? I promise you'll love me! ???:*??? I am 100% independent.....***-***-4827

December 2014:
•:*• Hi, I'm Nikki•:*• •:*• For those guys who expect the BEST, stop searching My services are of TOP quality and I can guarantee satisfaction. •:*• •:*•Take a break from your stressful lifestyle and enjoy a fun-filled relaxing time with me.... •:*• I promise you'll love me! •:*• I am 100% independent!! 100%REAL PICS!!! Call Me (***) ***-9983

Name and Phone



Hi guys my name is Mandy and im all real. If you are not getting your needs taken care of at home come see me. Im not like the rest on here. With me its all fun and no games or drama. Im 5 2 145lbs 36D long hair and tan. Safe place not a motel or hotel and desceate. Im work for myself and dont do drugs. So what are you waiting for call me and lets play. *** *** 9151. please be respectful. Price includes all. Overnites avaiable. NO AFERICAN AMAERICAN MEN

634 days

Hey guys i'ts Mandy and its my last weekend here i'm moving back home. I'm 5 2 145 38D long black hair. Toes are done for you foot fetish men. Pictures are 100% me and yes guys I will be wearing the outfit in the pictures tonight. Come get while you can. 412-pls no black men. *** ***-8568

427 days

Hey guys Come enjoy my great skills. There is no bs and no drama just a good a time.So come play with me and let me show you what i can do. Im independent and am alone. With me you dont have to worry. Please no black guys and blocked numbers will not be answered. Up all night so lets get the party started. ***-***-8568. Leaving in the am. In a nice clean place not a run down motel. All real pics.

Future Work

- Incorporation of new similarity measures
- Further leverage visual information using background and foreground segmentation and matching

References

- [1] I. L. Organization, "Global estimates of modern slavery: Forced labour and forced marriage," 2017.
- [2] M. Latonero, "Human trafficking online: The role of social networking sites and online classifieds," Available at SSRN 2045851, 2011.

Acknowledgements

I would like to thank Anthony Wertz and Dr. Artur Dubrawski, and the rest of the Auton Lab for their advice. I would also like to thank Rachel Burcin and Dr. Dolan for all their support. Finally, I would like to thank the RISS cohort for a wonderful summer and shared memories.

RISS ROBOTICS INSTITUTE SUMMER SCHOLARS
THE ROBOTICS INSTITUTE
Auton Lab
NSF