



Avoid Crowding in the Battlefield: Semantic Placement of Social Messages in Entertainment Programs

(work in progress!)

TCS Research

AI4TV Workshop, ACM Multimedia

October 2020



Authors

Yashaswi Rauthan, Vatsala Singh, Rishabh Agrawal, Satej Kadlay,
Niranjan Pedanekar, Shirish Karande, Manasi Malik, Iaphi Tariang

Picture from <https://topperababhas.wordpress.com/2014/07/08/be-like-a-train/>

Social Messages



Images: *Guide* (Navketan films), World Health Organization

Wash your hands

Wash your hands with soap and running water when **hands are visibly dirty**



If your **hands are not visibly dirty**, frequently clean them by using alcohol-based hand rub or soap and water



Social Messaging - Challenges

1

Messages need to reach a **large** population with varying **demographics**

2

Messages get **drowned** in social media

3

Messages need to be **controlled** by official entities and **localized** if required

4

Messages need to have **familiarity** and **recall**



77 million viewers

Placing Messages in Entertainment Content

Entertainment-Education Model

Educational messaging 'sugar-coated' in entertainment lowers audience resistance and defends the embedded messages

Arvind Singhal and Everett Rogers. 2012. Entertainment-education: A communication strategy for social change.

Till now, social messaging has been embedded in the content

Can it be done programmatically?



Extended Elaboration Likelihood Model

Viewers are less likely to counter-argue messaging that is embedded in entertainment content, since they are engaged in the dramatic elements of the content

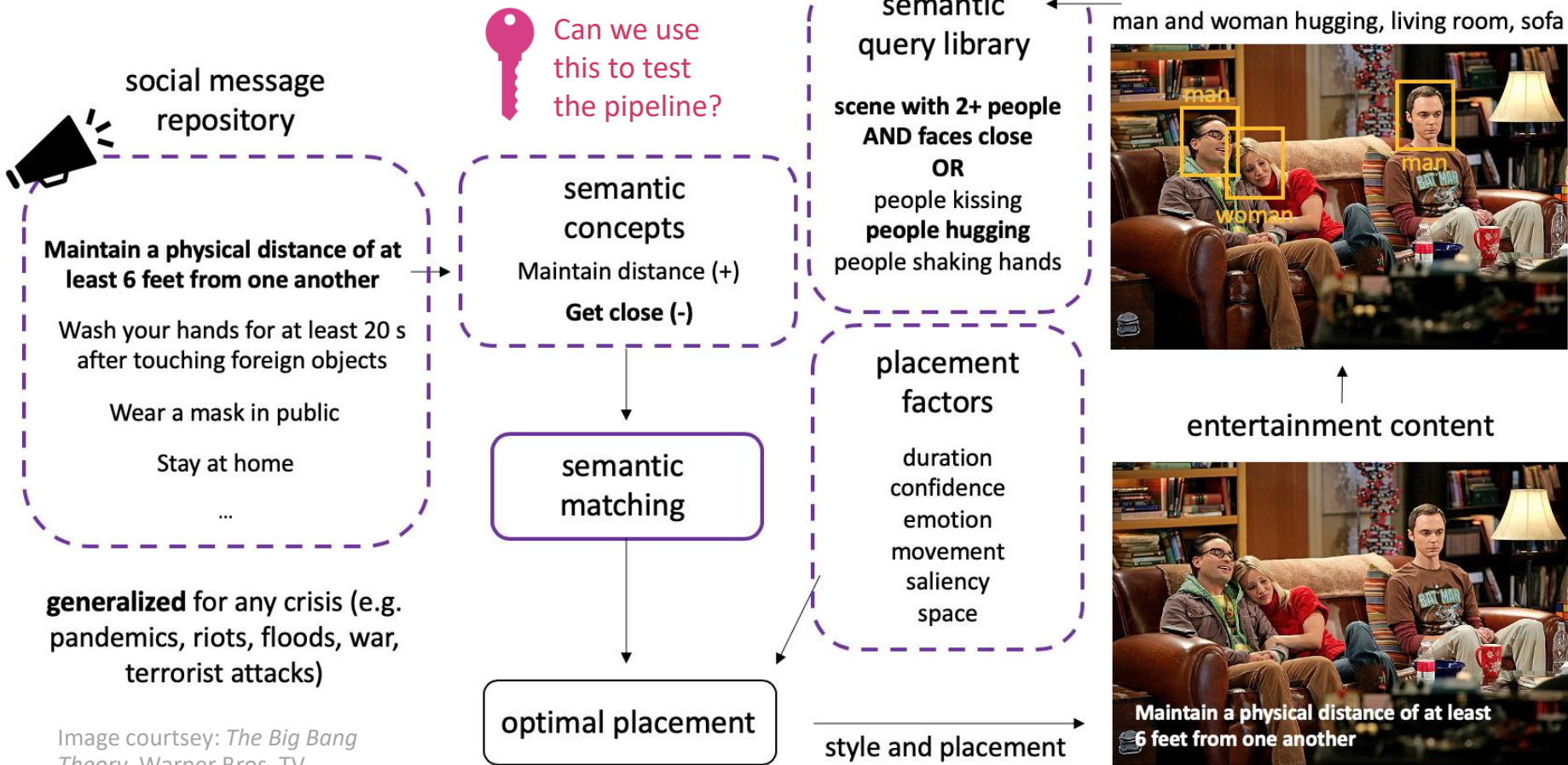
Michael D Slater and Donna Rouner. Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. Communication theory 12, 2 (2002), 173–191.

Contextual advertising

Tends to have better recall with audiences

Yadati et al. CAVVA: Computational affective video-in-video advertising. IEEE Transactions on Multimedia (2014).

Proposed Method



Messages and Behaviours

COVID 19 Crisis

Message ID Social Message	+ve Behavior	-ve Behavior
M1 <i>Stay at home as far as possible and avoid going to crowded places</i>	Stay at home	Go out, Be seen in crowds
M2 <i>Wash your hands for at least 20 seconds after touching foreign objects</i>	Wash your hands	Touch foreign objects
M3 <i>Maintain a physical distance of at least 6 feet from one another</i>	Stay far apart	Get close
M4 <i>Wear a mask while going out</i>	Cover your face	Be seen out without a mask
M5 <i>Avoid touching eyes, nose and mouth</i>	Keep hands off the face	Touch eyes, nose or mouth

Terrorism Threat

Message ID Social Message	+ve Behavior	-ve Behavior
M6 <i>Be extra vigilant in crowded places</i>	Look around	Be on your own, watch mobile, read
M7 <i>Do not leave your belongings unattended</i>	Carry belongings (sacks, bags)	Stay away from belongings
M8 <i>Do not touch unidentified packages</i>	Stay away from packages	Touch a package

Media Content

3 episodes of
Kuch Rang Pyar Ke Aise Bhi (daily soap serial)



Image courtesy: Sony TV



Crash (modern drama)

Image courtesy: Lionsgate Films,
DreamWorks Pictures, New Line Cinema.



Gladiator (historical drama)



Lord of the Rings (fantasy drama)

Annotations

Shot Boundaries (*Lokoč et al. 2019*)

Faces (*FaceNet, Schroff et al. 2015*)

Objects (*YOLO3, Redmon & Farhadi, 2018*)

Persons (from *YOLO3*)

Places (*Places, Zou et al. 2015*)

Activities (*I3D, Carreira & Zisserman, 2017*)

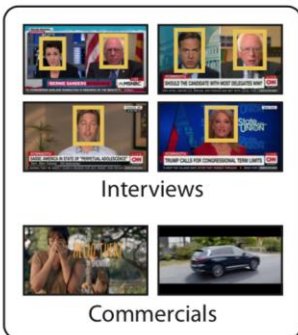
Body Parts (*Neverova et al. 2019*)

Expressing semantics in terms of annotations

Behavior	Query Structure
Go out	person=true AND place=outside
Be seen in crowds	count(person)>n AND place=outside
Touch foreign objects	object=true AND person=true AND body_part=hand NEAR face=true
Get close	count(person)=2 AND (<i>face</i> ₁ NEAR <i>face</i> ₂ OR activity=kissing, hugging, handshake)
Go out without a mask	person=true AND place=outside
Touch face	face=true AND body_part=hand NEAR face

Semantic Querying

```
def bernie_and_jake(faces):  
    bernie = faces  
        .filter(face.name == "Bernie")  
    jake = faces  
        .filter(face.name == "Jake")  
  
    bernie_and_jake = bernie  
        .join(jake,  
            predicate = time_overlaps,  
            merge_op = span)  
  
    return bernie_and_jake
```



*Fu et al., Rekall: Specifying Video Events using Compositions of Spatiotemporal Labels (2019).

Queries using Rekall*

Spatiotemporal query specifications for videos

Coalesce over smaller segments

For in-between frames not returned by queries

a representative query

Q1 Find frames with: $\text{count}(\text{person}) \geq 2$ AND $\min(\text{dist}(\text{centroid}(\text{face}_i), \text{centroid}(\text{face}_j))) \leq 0.1 * \text{width}(\text{frame})$

Q2 Find frames with: (activity="kissing" OR activity = "hugging" OR activity = "handshake")

a candidate location

union of all query results with a coalesce interval
returns start frame, end frame, average confidence

Multiple queries for the same message (10 for 5)

Using multiple annotations

Confidence measure using heuristics

Weighted mean of binary output of queries per frame basis
based on robustness of annotators

Optimal Placement

Input

Each candidate location with confidence and duration

Constraints

A time slot is an equal division of content duration in n parts (n = number of messages)

One message at the most for one time slot, though a message need not have an assignment

Treated as an **assignment** problem with a cost matrix using a product of normalized confidence and duration (using *OR Tools*)

a representative cost matrix

	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5
Msg 1	10	90	128	135	542
Msg 2	10	90	128	135	542
Msg 3	1000	10000	372	399	10000
Msg 4	460	450	310	462	300
Msg 5	744	668	744	200	706

Lord of the Rings: Return of the King (New Line Cinema)



Results – M3

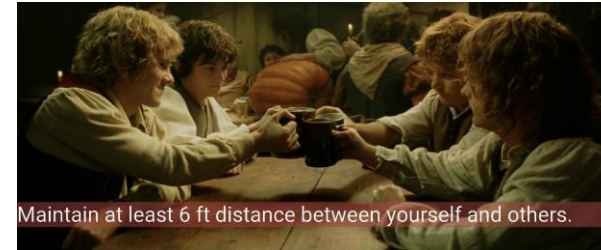


Image courtesy: Sony TV, Lionsgate Films, DreamWorks Pictures, New Line Cinema.

Results – M4

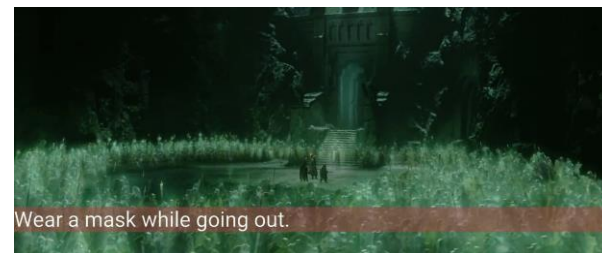


Image courtesy: Sony TV, Lionsgate Films, DreamWorks Pictures, New Line Cinema.

Results – M1

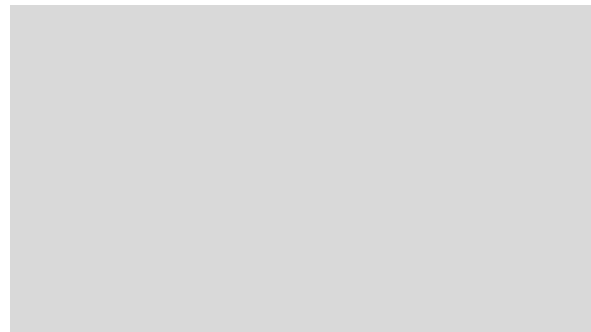
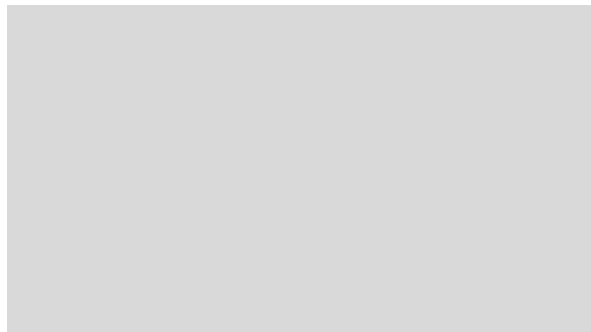


Image courtesy: Sony TV, Lionsgate Films, DreamWorks Pictures, New Line Cinema.

Results – M2



✗



✗



✗



✓



✓



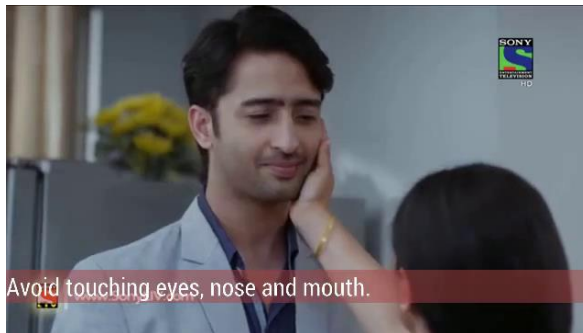
✓

Image courtesy: Lionsgate Films, DreamWorks Pictures, New Line Cinema.

Results – M5



✗



?



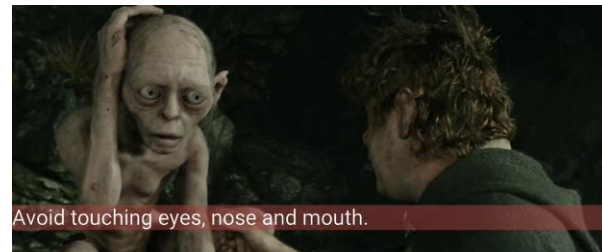
✓



✗



?



✗

Image courtesy: Sony TV, Lionsgate Films, DreamWorks Pictures, New Line Cinema.

Results Analysis

Message ID Social Message	CC	C/h	Dur	Conf	CP
<i>M1 Stay at home as far as possible and avoid going to crowded places</i>	283	17.5	1.54	0.73	0.982
<i>M2 Wash your hands for at least 20 seconds after touching foreign objects</i>	18	2.5	2.13	0.59	0.594
<i>M3 Maintain a physical distance of at least 6 feet from one another</i>	1183	115.5	1.42	0.50	0.955
<i>M4 Wear a mask while going out</i>	283	17.5	1.54	0.73	0.971
<i>M5 Avoid touching eyes, nose and mouth</i>	105	11.4	0.83	0.80	0.574

CC

Candidate Count

C/h

Candidates per hour of media

Dur

Median duration of candidates

Conf

Median confidence of candidates

CP

Candidate precision

Some events are more frequent than others

High precision for some queries

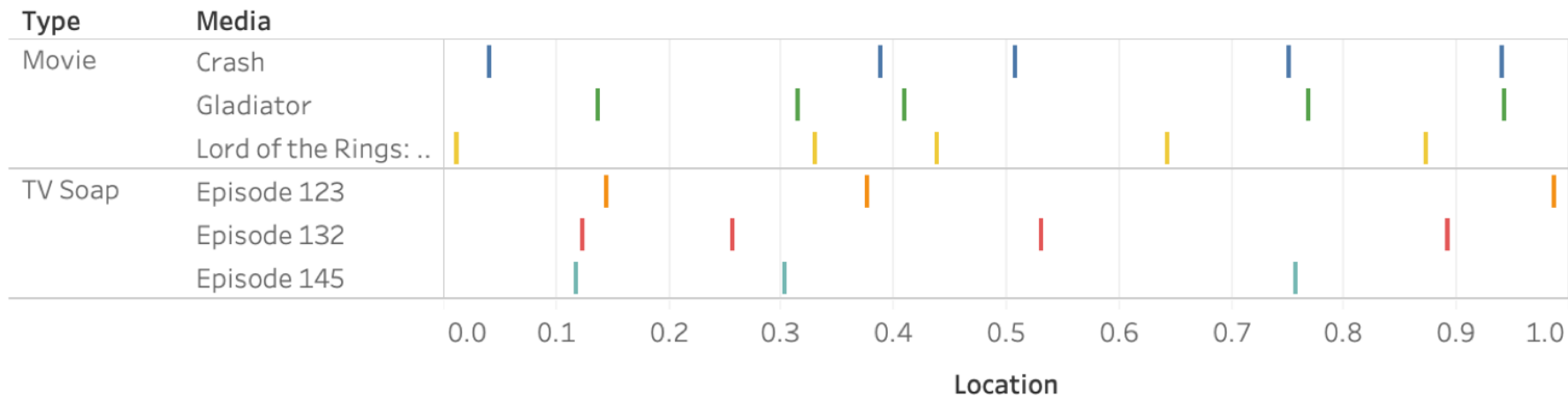
Base annotations are less noisy (e.g. persons, places)

Low precision for some queries

Base annotations are noisier (e.g. activity, body parts)

Confidence measure needs tweaking

Spacing between Messages



Conclusions

Spatio-temporal querying can work

Need survey-based results and wider content

Juxtaposition produces interesting results

Unlike the 'smoking kills' message

How to learn queries

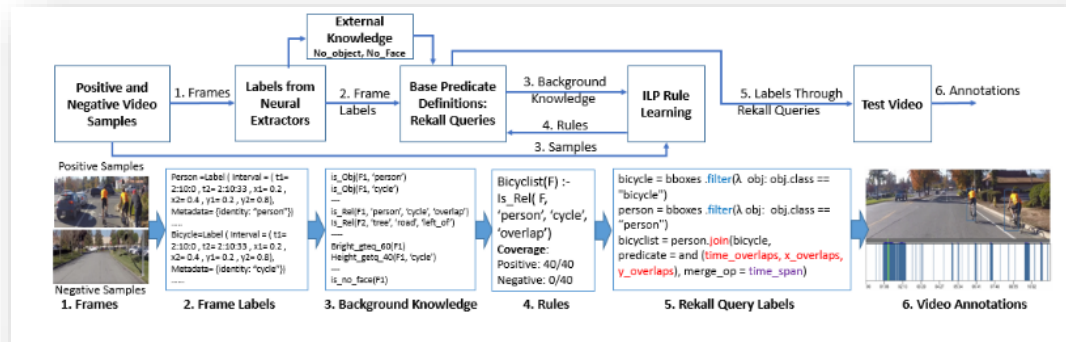
Query building is manual and takes effort

Learning queries using spatio-temporal algebra?

Need to use affective annotations

Negative and positive placements

To avoid wrong placements



Sivaprasad et al. Multimodal Continuous Prediction of Emotions in Movies using Long Short-Term Memory Networks. ICMR 2018.

Sivaprasad et al. Partners in Crime: Utilizing Arousal-Valence Relationship for Continuous Prediction of Valence in Movies. AAAI AffCon, 2019.

Next Steps

A bigger survey

A bigger variety of messages, content and viewers

A platform for social messaging in crisis viable?

A message repository and a query library

Difficult to adopt this with high-end content

Can this be taken to official TV networks? (e.g. DD)

What can replace messages?

Ad placement

Cross-media recommendations

Personal reminders (integration with calendar / Alexa)

Targeted physical / mental health messages

