

The First Social Customer Relationship Management System to Analyze Large On-line Social Networks

Tzu-Chi Yen, Xiaodong Liu[†] and Wuyang (Tony) Zhao[†]

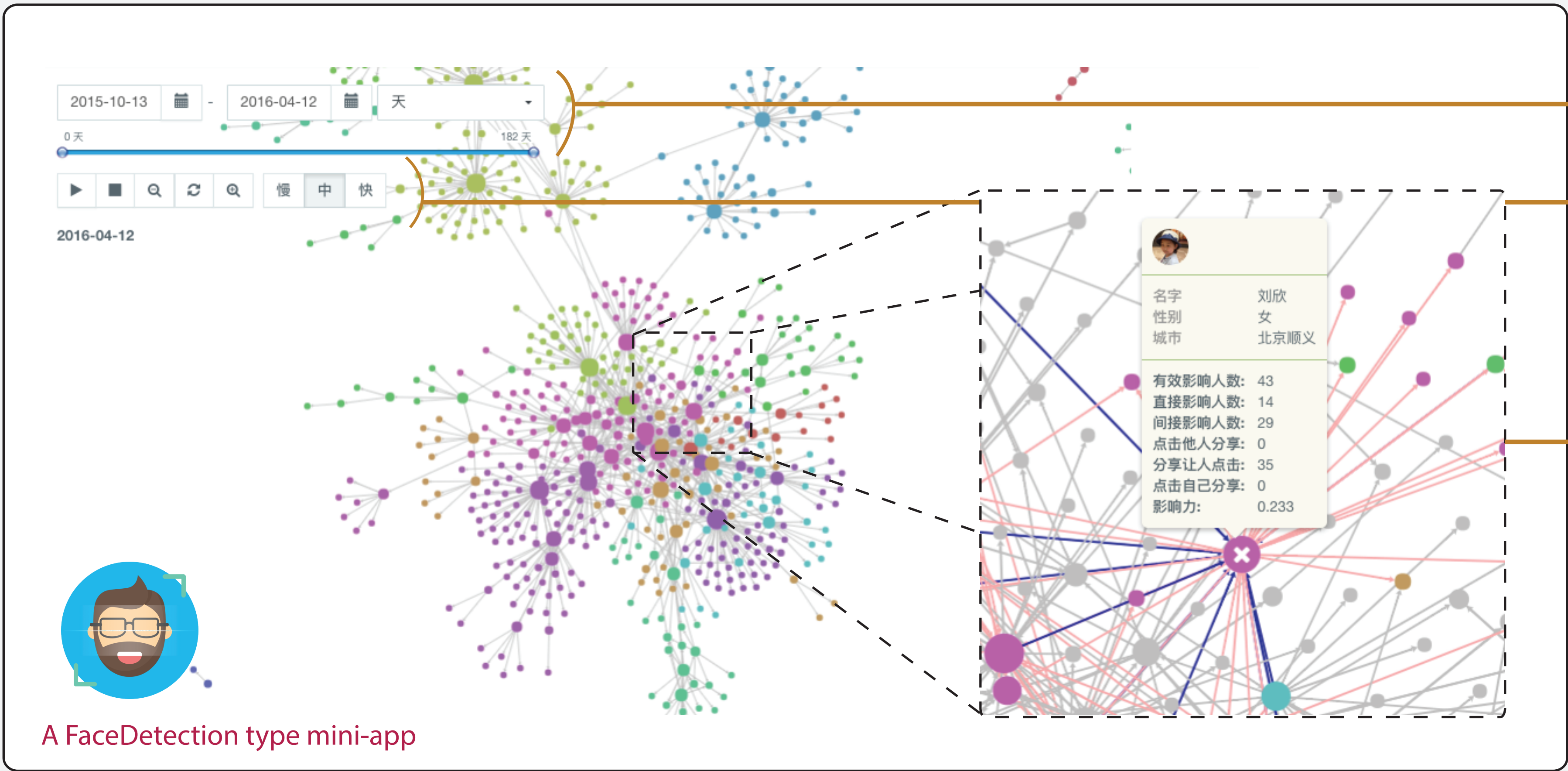
[†] Sensoro Technology Co., Ltd., Beijing City, China 100102

Abstract

We present a novel social customer relationship management system (Social CRM) that can collect, visualize, and analyze a customer network [1, 2]. Its working procedures are, (1) it provides various HTML5 mini-app templates for clients to customize into their marketing campaign which can only be accessed through the WeChat mobile app; (2) it reaches the clients' customers when they visit the shops to scan a QR code with their phone or click the post that derived from the mini-app which his/her friends shared on the WeChat Moments Wall; (3) when a customer, A, accesses the page posted by another customer, B, we assume that the campaign reaches A from B; (4) we retrieve the logs of share events containing a source/target OpenId, and a timestamp through the WeChat developer APIs; (5) we reconstruct the spreading process as a network using the logs; (6) finally, we visualize and identify influencers in the network for targeted marketing. In the figure of system overview below, we show a typical network. Each node represents an individual that had accessed the mini-app, and the directed edges associated with the node indicate who else were involved in the process.

The system is now deployed in Sensoro [3], with clients mostly running a restaurant. It has attracted 18,000,000 independent customer accesses since Dec. 8, 2015, and collected hundreds of independent networks, one for a spreading process, with the largest one containing 88211 nodes and 64758 edges. This is the first social CRM to our knowledge as a social analytic engine [4] for large on-line networks.

System Overview – the Network Analysis Web Module



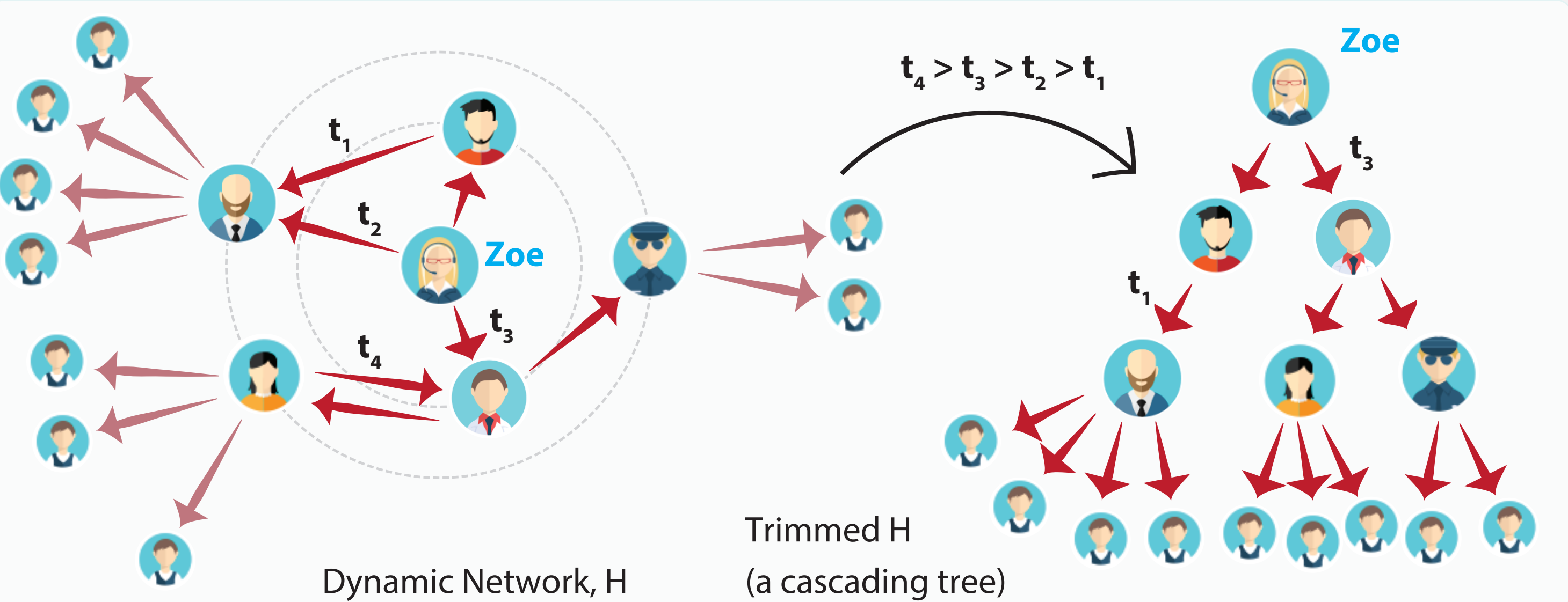
- Time panel and scroll bar: segmenting timestamps into hours or days
– One can time snapshot the network at ease
- Control toggles: zoom in/out, turning on the evolution of network with three paces
– Complex network evolution can be **directly seen from the browser**
- Infowindow (with clickable nodes), showing the node-type attributes, such as:
 - avatar
 - nickname, gender, city of living
 - measures of spreading influence
- When a node is clicked, the two colors of its belonging edges show the inbound links (blue, when the node clicks its friends), and the outbound links (pink, when other nodes click the node)
- The graph layout is pre-calculated as **SFDP format**, and the communities (no labelling, only pure algorithmic outcome) is carried out in the **louvain** way.

Message Spreading Process as a Network

We reconstruct the spreading process as a dynamic network using individual logs from WeChat's developer APIs. Each log represent a 3-tuple, (u, v, t) , that denotes a directed share event from user u to v at time t , where each user is characterized by an unique OpenId for the particular client.

Then, we assemble the network, $H = (V, E_t, X)$, where V is a set of users, $E_t \subseteq V \times V$ is a set of links between users, $e_{u,v,t} \in E_t$, denotes a directed link from user u to v , $(u, v) \in V$, at time t , and X is a $|V| \times K$ attribute matrix that contains other knowledge about the user, such as his/her gender, city of living, interests, etc.

Note H is embedded in a larger, possibly unknown, static social network, G .



We define four measures for each node in the dynamic network, H , to identify and rank the influential customers for targeted marketing.

Name of Measure	Property	Description	Zoe's Score
Outdegree	Structural	Number of outward edges	3
Betweenness	Structural	Normalized directed centrality	1
Effective Outdegree	Phenomenological	Outdeg. in the cascading tree	2
Effective Influence	Phenomenological	Number of downstream nodes in the cascading tree	14

* via depth-first search

Some Open Research Problems

Question 1: How to learn a model and parameters that fit well the spreading process in the data?

Question 2: How to utilize the rich node-type metadata to contribute the "attribute-based" influence maximization problem? For example, how to find seeds to deliver memes that influences the most people who live in Seoul?

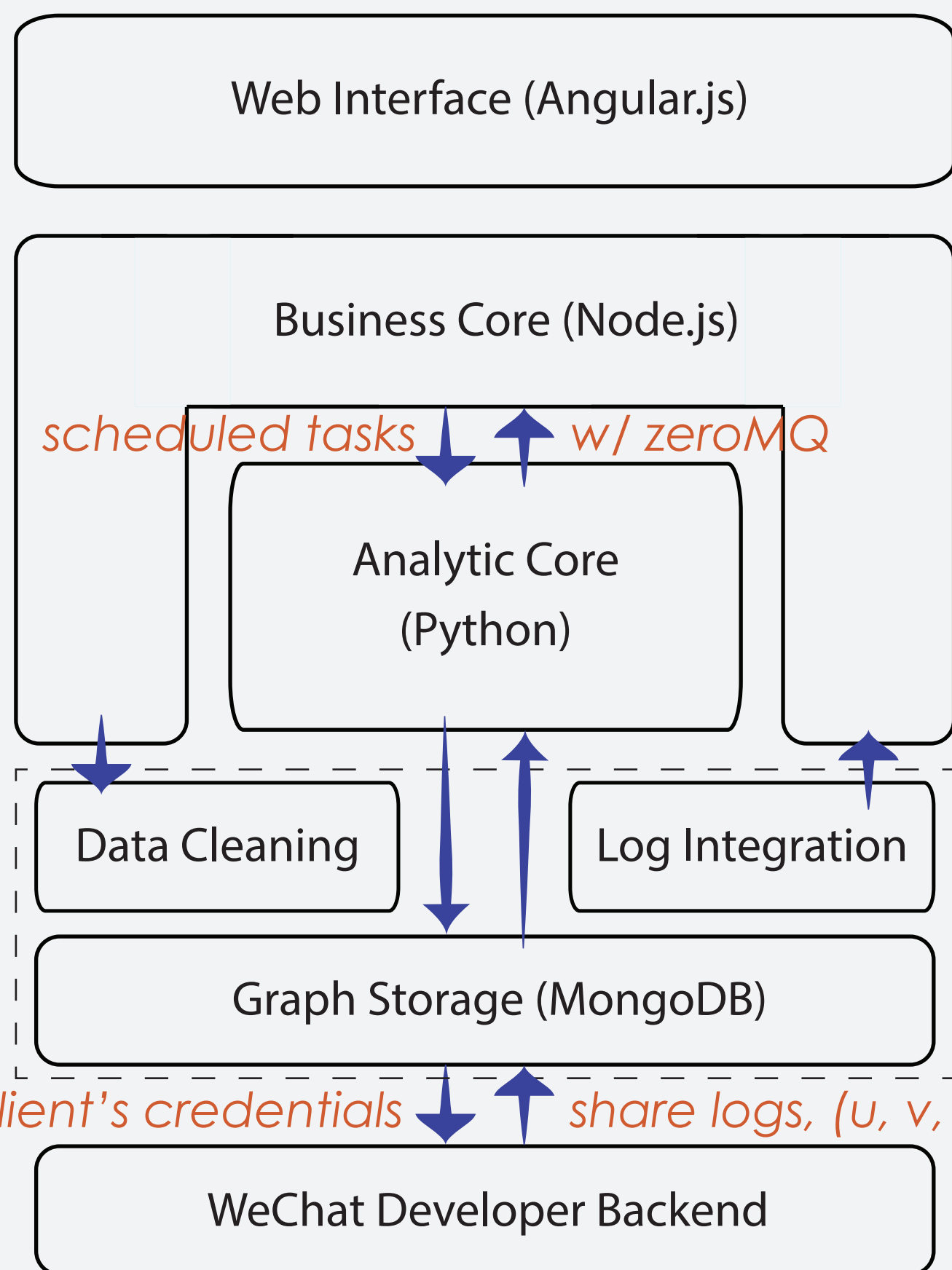
Question 3: Can we predict the popularity of a marketing campaign?

References

- [1] Sensoro's Social CRM Website (Commercial System): <http://x.sensoro.com>
- [2] The DEMO video (1 min): <http://junipertcy.info.s3.amazonaws.com/netsci-16.m4v>
- [3] Sensoro is a iBeacon/IoT startup in Beijing: <http://www.sensoro.com>
- [4] Y. Yang, et al., SAE: social analytic engine for large networks, in KDD'13, pp. 1502-1505, 2013.

System Infrastructure

We create the Social CRM system using the MongoDB–ExpressJS–AngularJS–NodeJS (MEAN) web programming stack, which communicates with an independent Python process that calls the graph-tool module (implemented in C++) for intensive calculations, such as centrality, graph layout, and graph partition.



– for $H = (V \sim 10^4, E \sim 10^4)$, total calculation time ~ 3 min [2.1 GHz, 4 GB RAM]

Name of app	#Networks	#Nodes*	#Edges*	Description (as a campaign)
Articles	854	18419 (16347)	19900 (19004)	<i>Hey, come read this interesting article!</i>
FaceDetection	530	5430 (4315)	6024 (5713)	<i>Upload your photo to get a free meal!</i>
XmasPlay	175	4411 (1114)	1587 (1326)	<i>Santa Claus hides the presents, where's it?</i>
SquaredUP	333	13709 (11744)	12648 (11841)	<i>Your luck turning around in the blocks!</i>
ScratchCard	469	84742 (52657)	63202 (60045)	<i>Scratch the card to get a discount.</i>
LotteryBox	633	11722 (686)	1886 (794)	<i>Share and let your friends help you!</i>

* largest network in the set; parenthesis for its largest component

Concluding Remarks

We present a Social CRM that is possibly the first one to store and analyse large on-line social networks through WeChat. The network is visualized dynamically directly from the web browser, with clickable nodes showing the metadata of each node. To promote targeted marketing, we calculate the effective influence measure for each node of the spreading network. We rank the nodes that our client can send them coupons directly from the Social CRM. Such bussiness rationale is proven effective in China.

The data is currently not publicly available. Please contact Tony <tony@sensoro.com>, CEO of Sensoro, if you would like to access the data. If you are just interested to work further on the topic (web system or network research), please contact Tzu-Chi <junipertcy@gmail.com> or tweet @oneofyen.

Acknowledgement

The presenter and corresponding author, Tzu-Chi, thanks Sensoro Technology Co., Ltd. for generously providing thier data and system for the NetSci2016 poster presentation.



QR code to the DEMO video