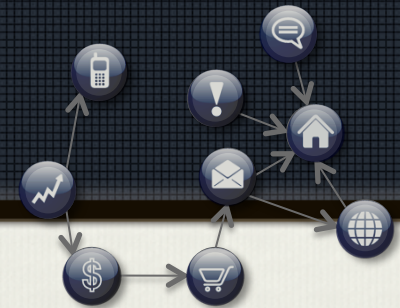


# NETWORKED USER ENGAGEMENT

USER ENGAGEMENT OPTIMIZATION WORKSHOP – CIKM 2013



***Janette Lehmann***

*In collaboration with:  
Mounia Lalmas,  
Ricardo Baeza-Yates,  
Elad Yom-Tov*

 UNIVERSITAT  
POMPEU FABRA

**YAHOO!**

# OUTLINE

## 1. Motivation

How to measure user engagement?

## 2. Network metrics

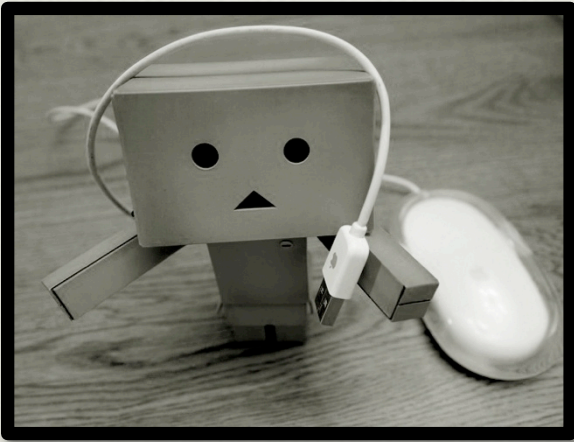
From site to network engagement.

## 3. Measuring networked user engagement

Case study based on the network of Yahoo sites.



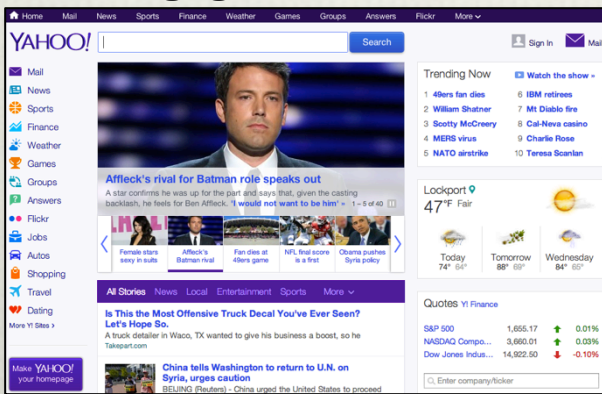
182-365+1 by meaganmikes



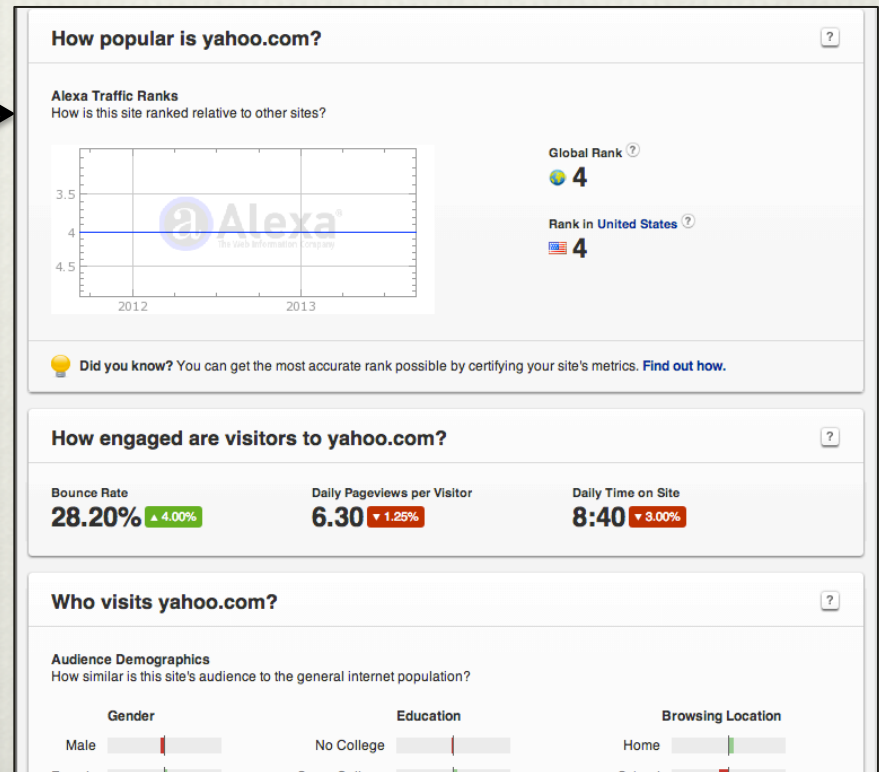
# How to measure User Engagement?

# MEASURING USER ENGAGEMENT

## User engagement on Yahoo!



## Web traffic reports: Google Analytics, Alexa.com, ...

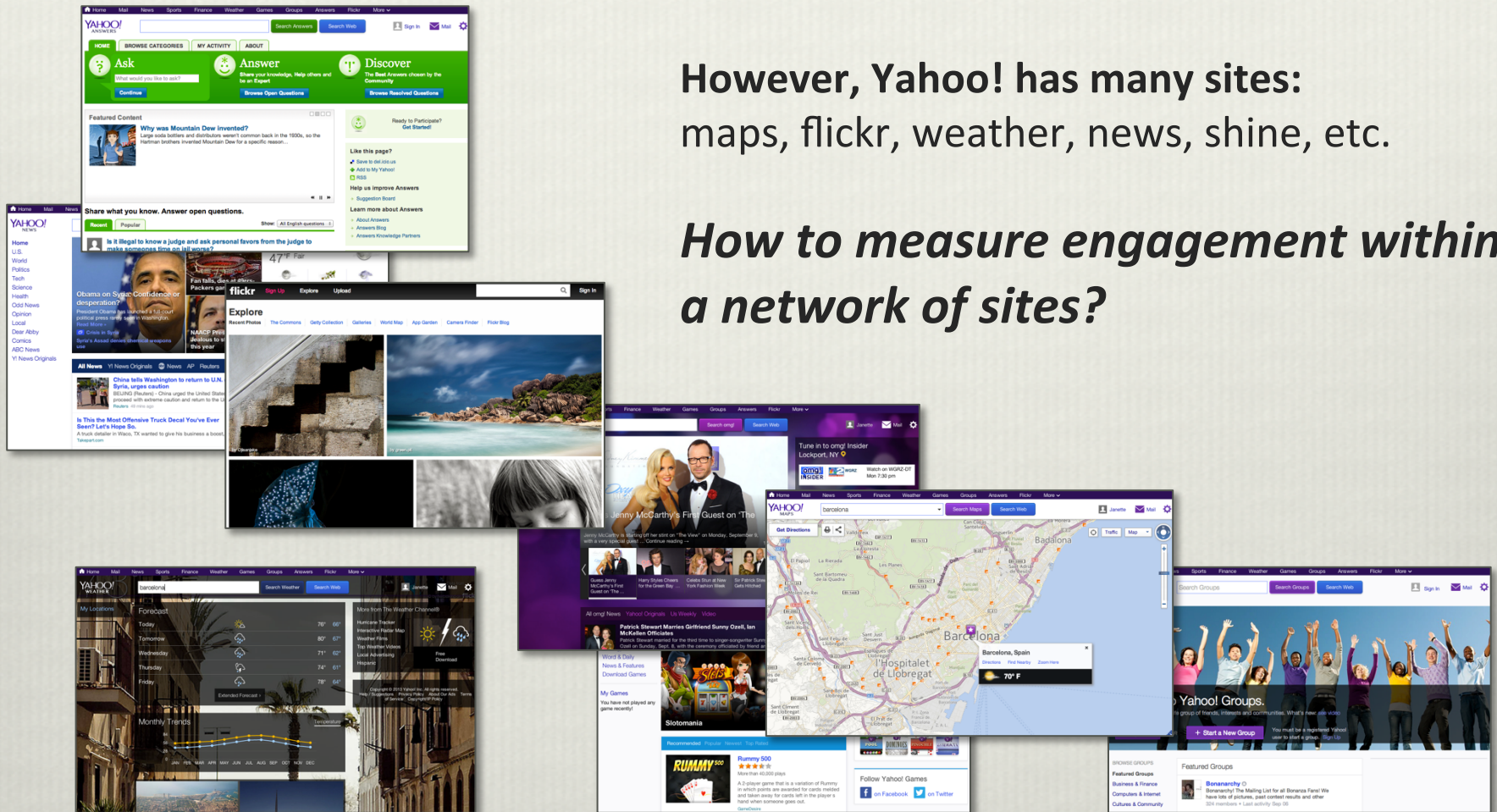


Popularity (#Users)  
 Activity (DwellTime)  
 Loyalty (ActiveDays)

# MEASURING USER ENGAGEMENT

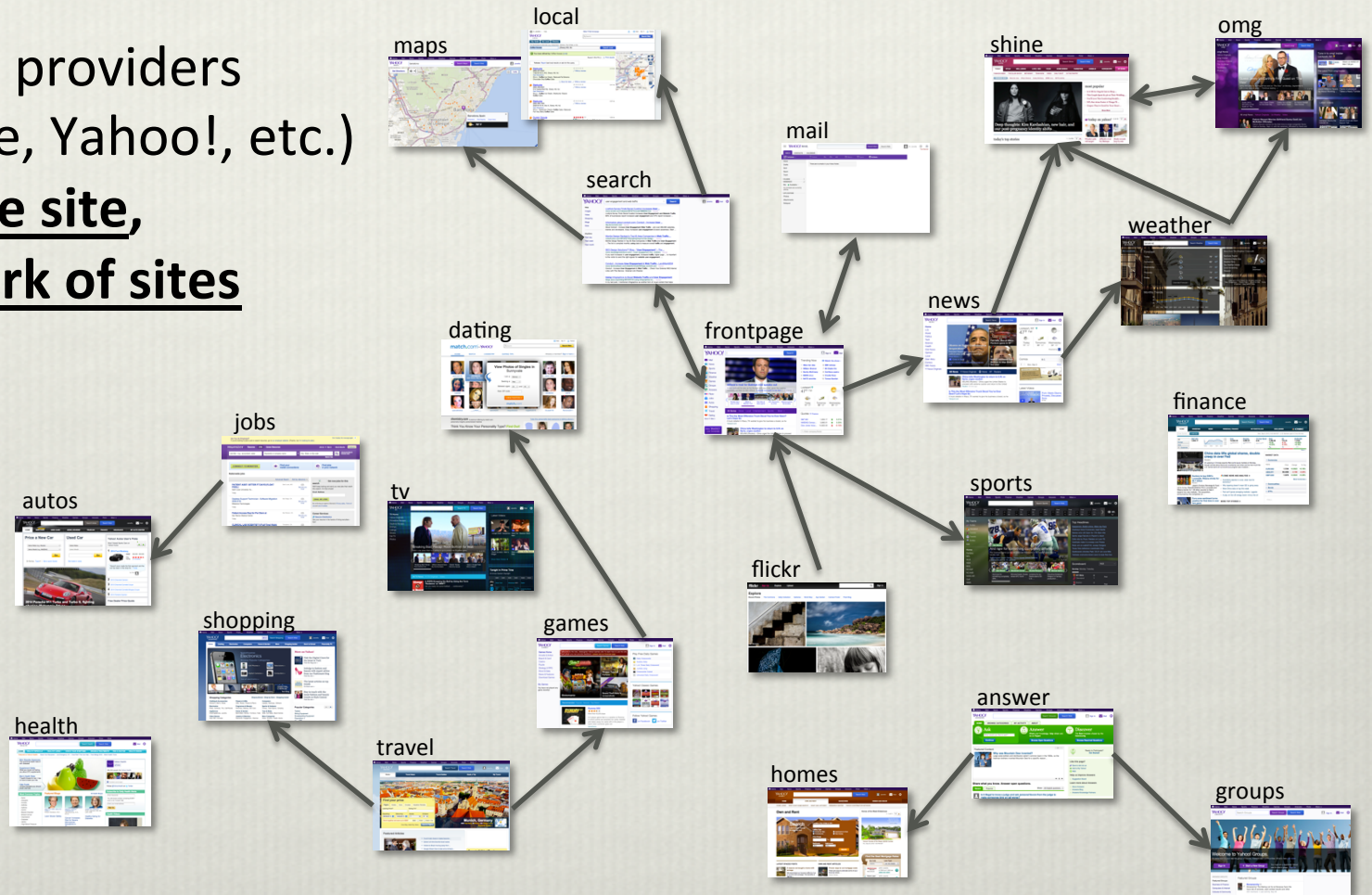
However, Yahoo! has many sites: maps, flickr, weather, news, shine, etc.

*How to measure engagement within a network of sites?*



# NETWORK OF SITES

Large online providers (AOL, Google, Yahoo!, etc.) offer not one site, but a network of sites



# NETWORK OF SITES

Each site is usually **optimized individually**, with some effort to direct users between them



# NETWORK OF SITES

## Online multitasking

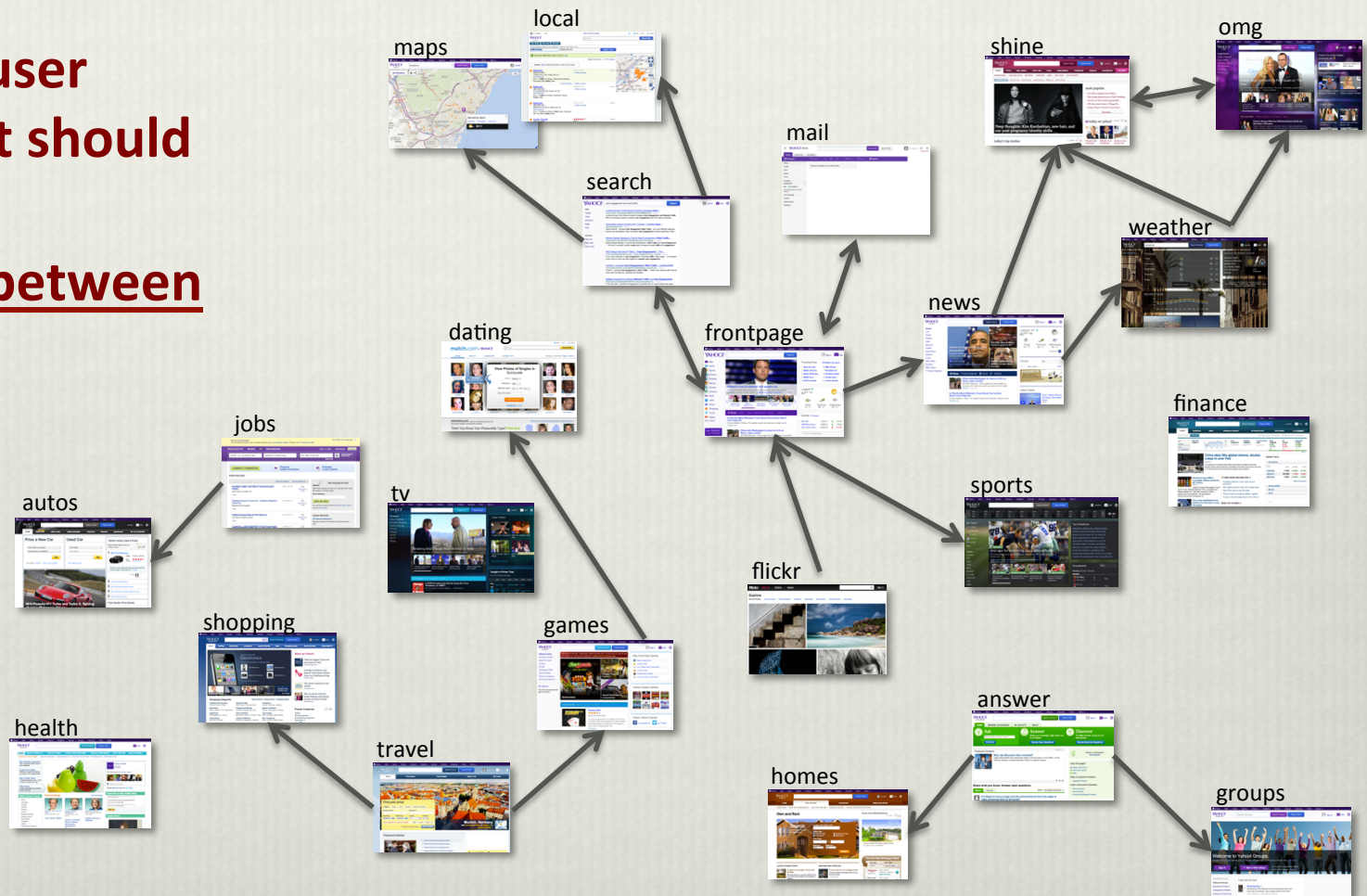
Users switch between sites within one online session (e.g. emailing, reading news, social networking, search)





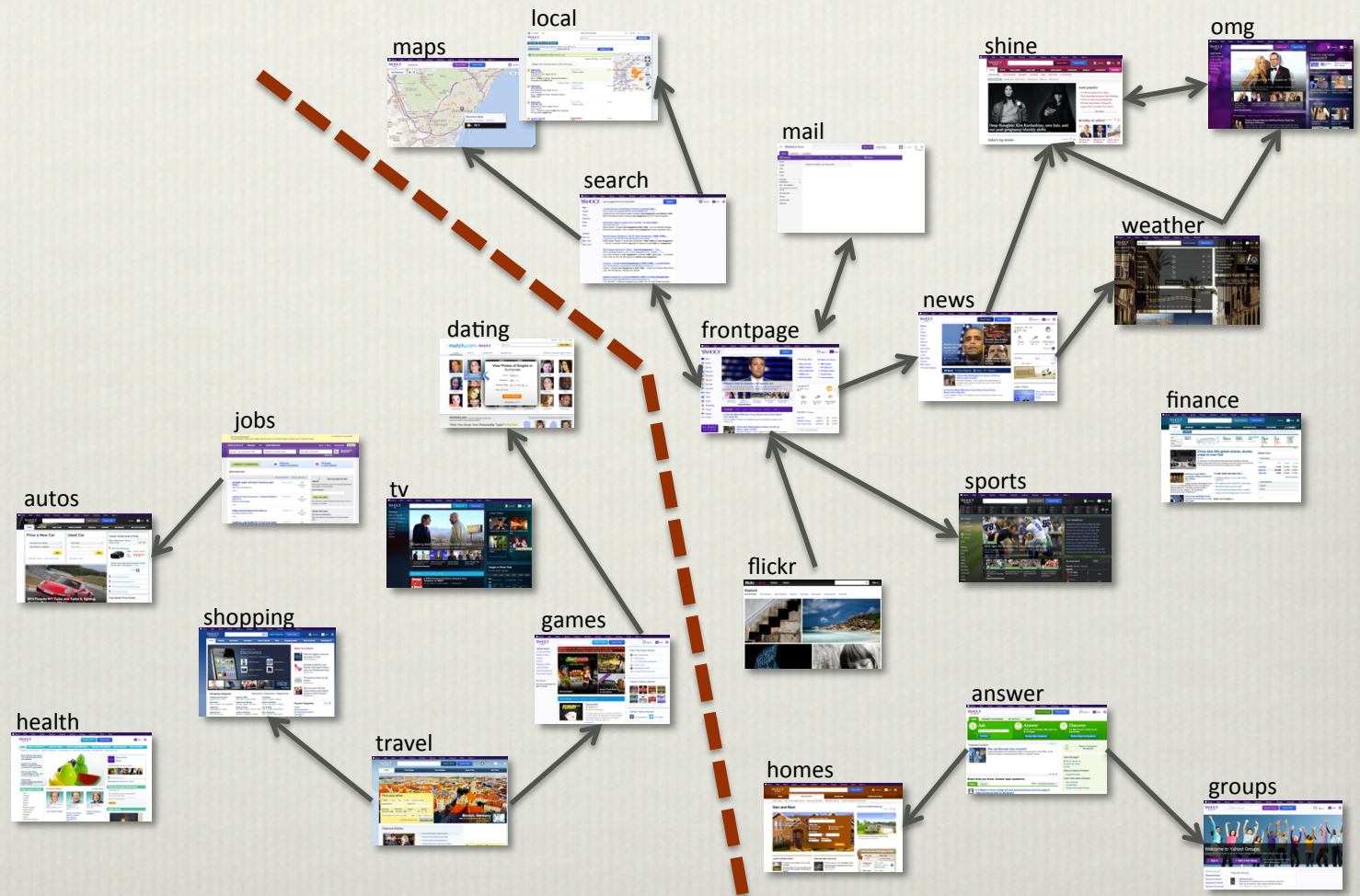
# NETWORK OF SITES

**Measuring user engagement should account for user traffic between sites...**



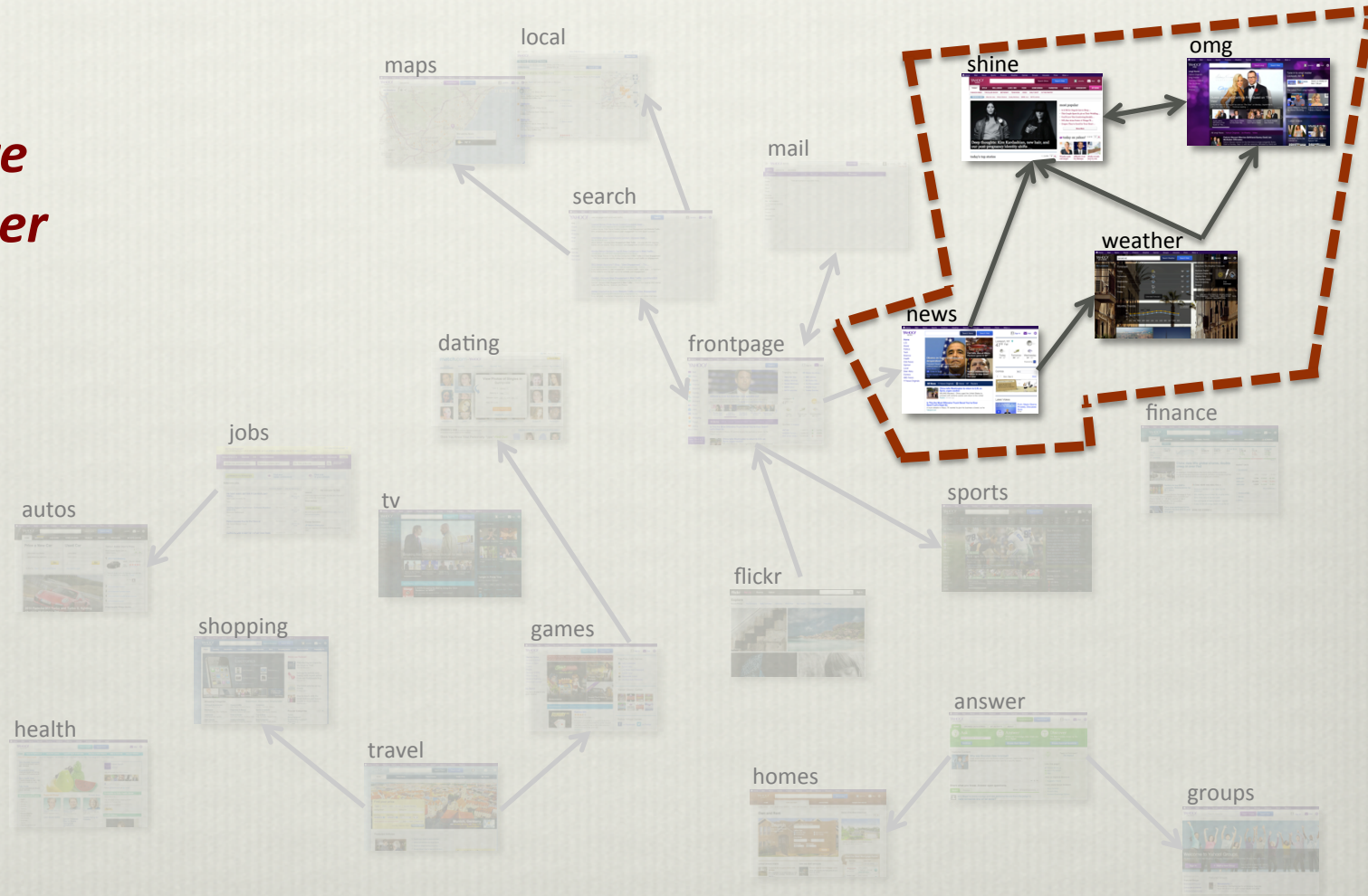
# NETWORK OF SITES

*... to detect missing connections*



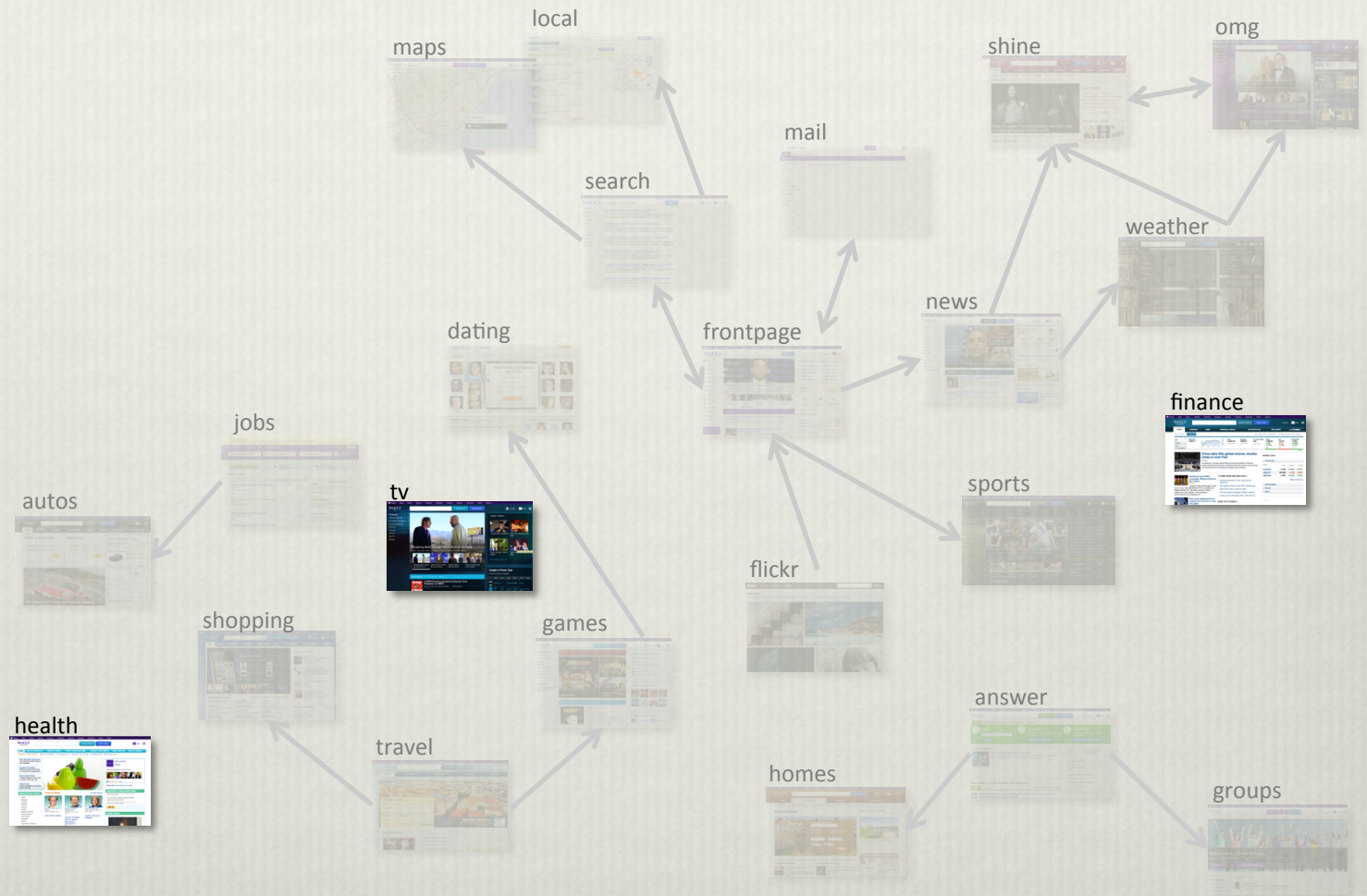
# NETWORK OF SITES

*... to detect sites that are used together*



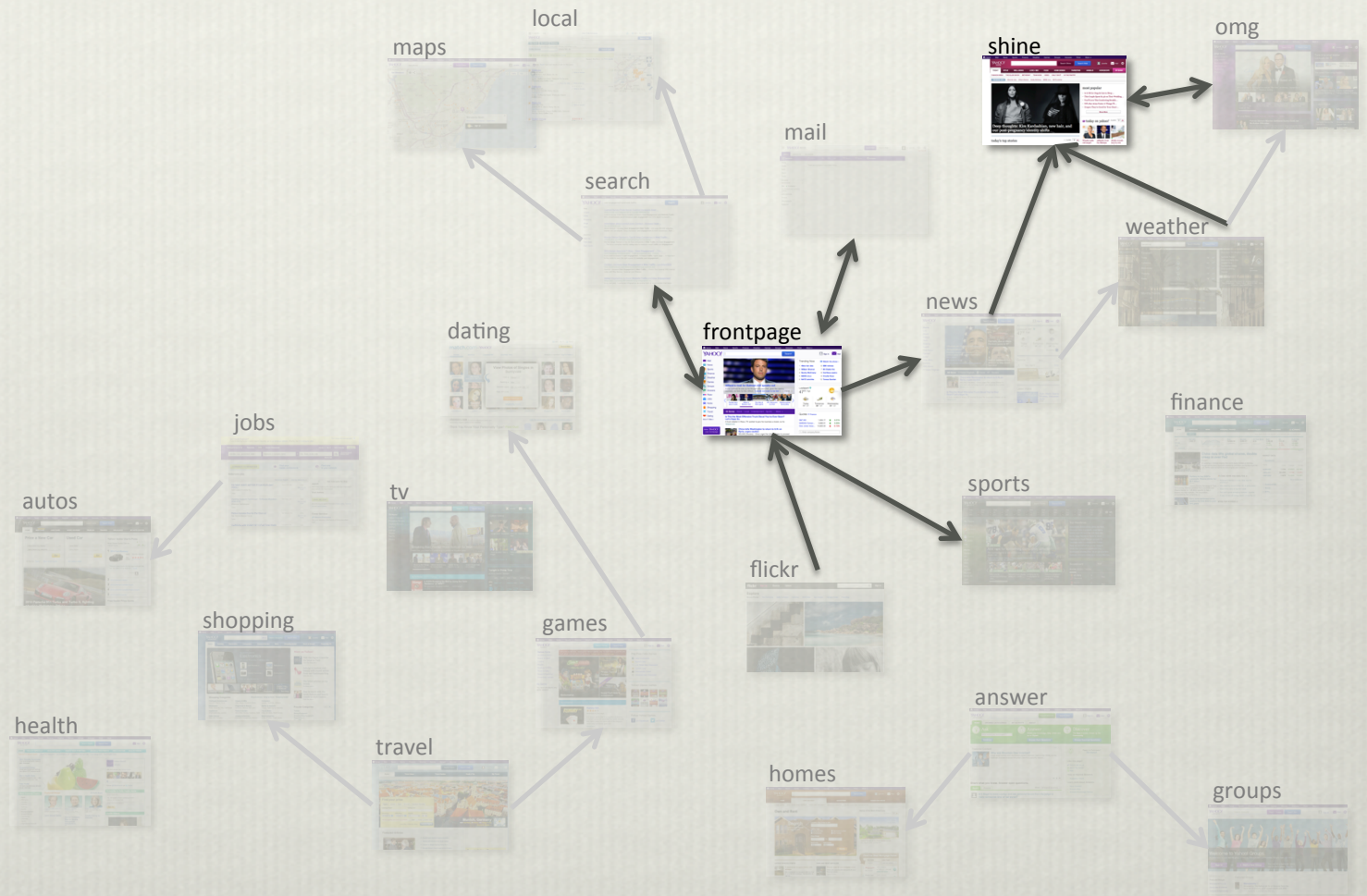
# NETWORK OF SITES

*... to detect isolated sites*



# NETWORK OF SITES

*... to detect important sites*

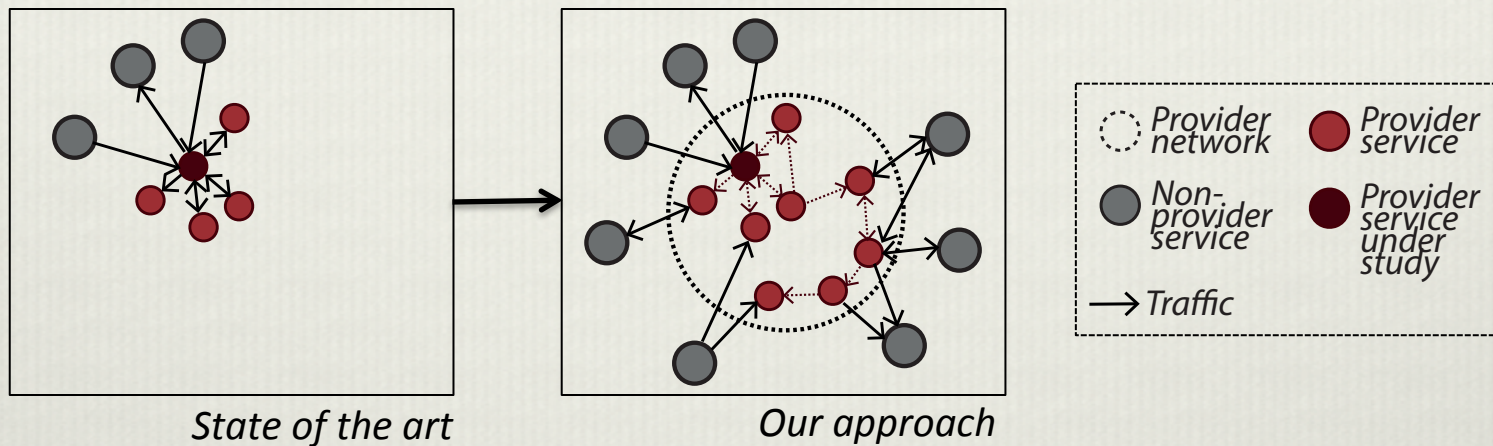




# From site to network engagement

Looking beyond one own nose!

# ENGAGEMENT NETWORKS



**$G=(V, E, \lambda)$**

V:  $n_{<i>}$  - Sites, services, functionalities, etc.  $n_{<e>}$  - External sites

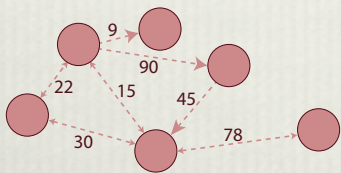
E:  $e_{<i>}$  - Traffic between internal nodes

$\lambda(e)$ : Traffic volume (#Users)

# NETWORK-LEVEL METRICS

## Network metrics

Size

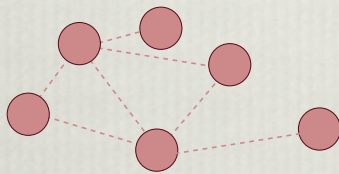


### **TrafficVolume**

Def.: Sum of edge weights

*A high value is a sign of a high networked user engagement; users navigate often between the sites of the network.*

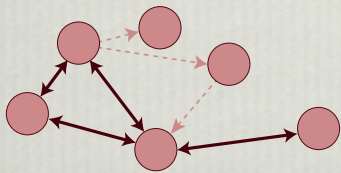
Connectivity



### **Density**

Def.: Proportion of edges to maximum possible

*A high value is a sign of a high networked user engagement; users navigate between many different sites.*



### **Reciprocity**

Def.: Ratio of number of edges in both directions

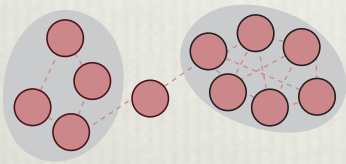
*A high value is a sign of a high networked user engagement; users do not only navigate from one site to another, they also tend to return to previously visited sites.*



# NETWORK-LEVEL METRICS

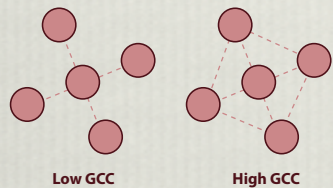
## Network metrics

Modularity



### **Modularity [ $\text{subNW}_{\text{mod}}$ ]**

Def.: Captures the existence of modules based on a random walk approach  
*A high modularity indicates that users visit many sites of one subnetwork, but hardly navigate to other subnetworks.*



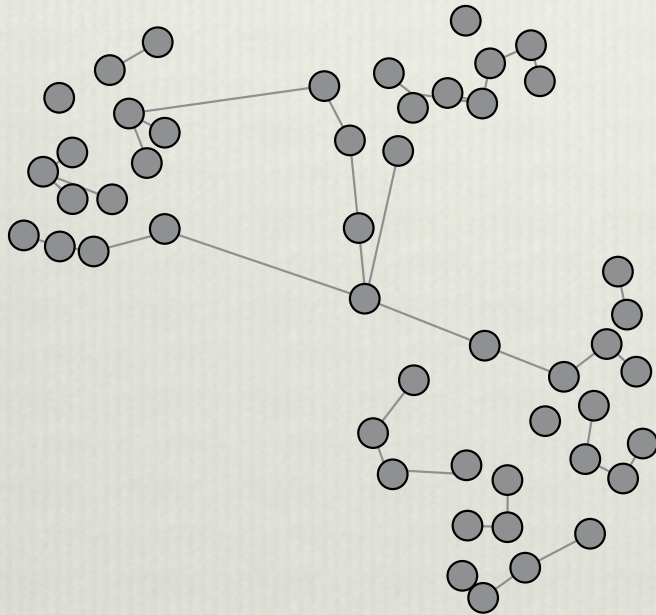
### **Global Clustering Coefficient [GCC]**

Def.: Captures the existence of tightly connected groups of nodes  
*A high value is a sign of a high networked user engagement; users access sites directly, instead of using front pages.*

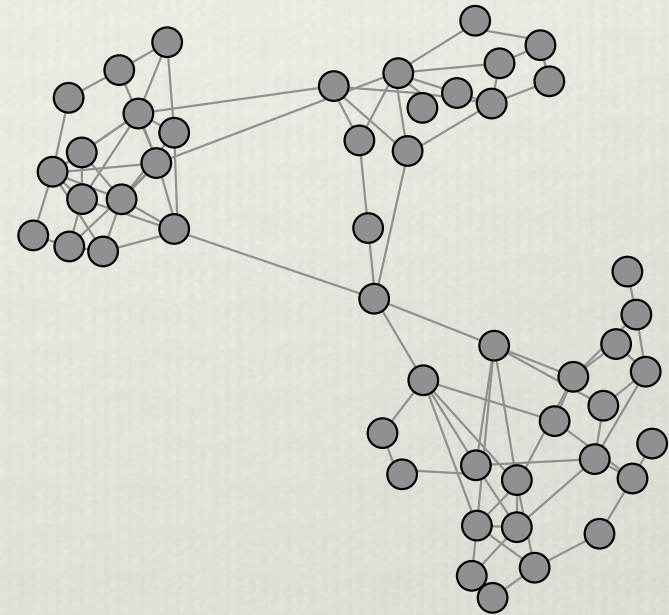
# ENGAGEMENT NETWORKS

## Modularity [ $\text{subNW}_{\text{mod}}$ ]

Yahoo! Network with 3 countries



*low modularity → low engagement*

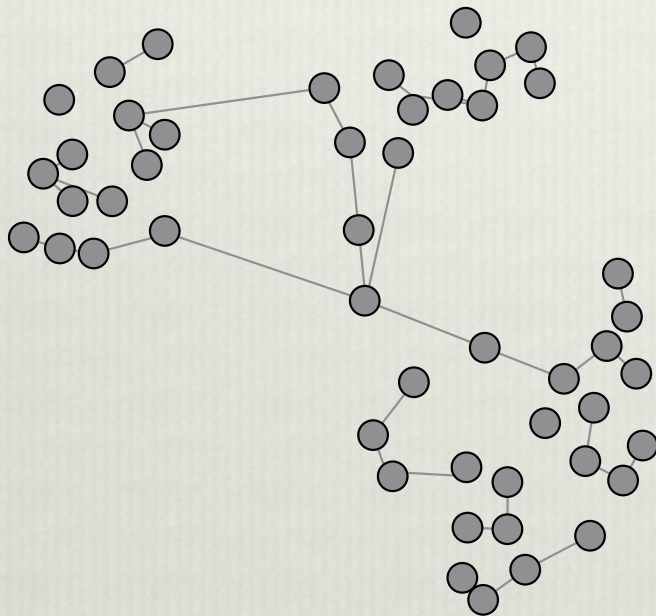


*high modularity → high engagement*

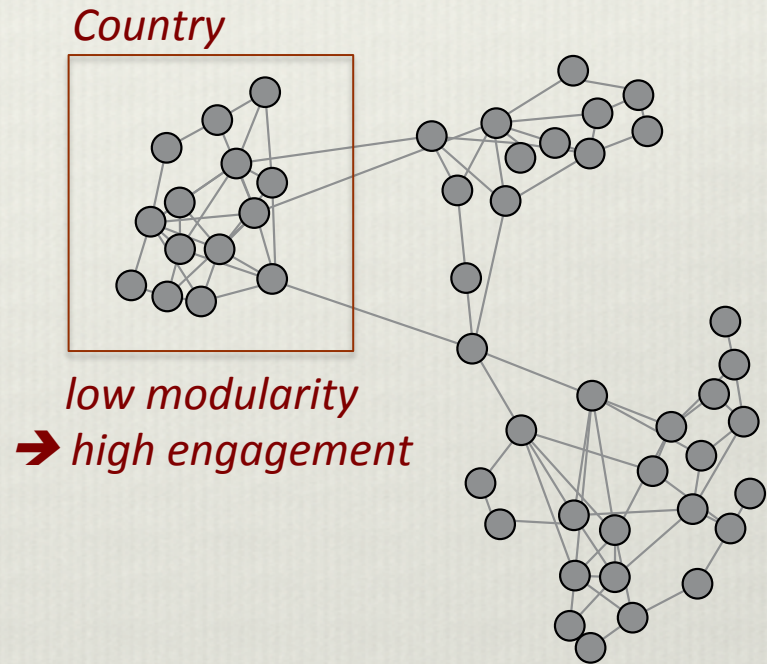
# ENGAGEMENT NETWORKS

## Modularity [ $\text{subNW}_{\text{mod}}$ ]

Yahoo! Network with 3 countries



*low modularity* → *low engagement*



*high modularity* → *high engagement*

left by [ embt ]



# Case Study

Measuring  
Networked user  
engagement

# ENGAGEMENT NETWORKS

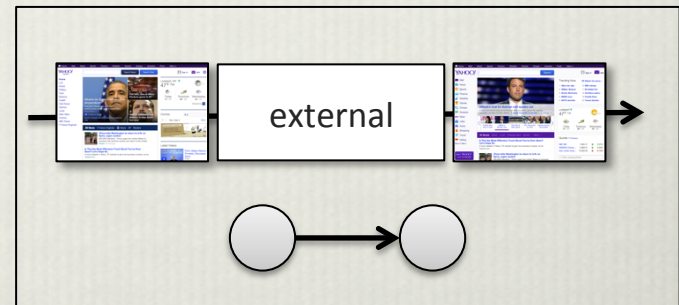
## Interaction data

- July 2011
- 2M user, 25M sessions
- 728 Yahoo! sites (news, mails, search, etc.)

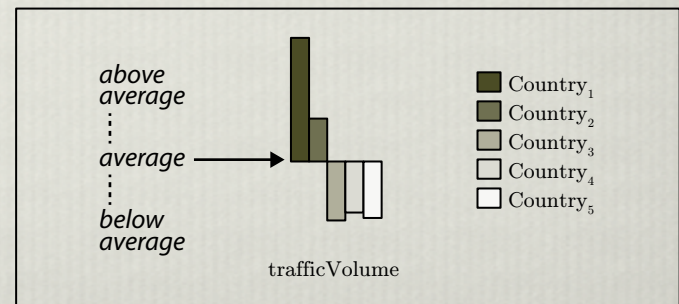
## Networks

- 12 weighted directed networks, filtered by
  - » Edge types (internal, internal + returning)
  - » Time (Wednesday, Sunday)
  - » User loyalty (causal, active, VIP)
  - » Country (5 EU countries)
- Applying metrics on networks, results are normalized using the z-score

*returning edge*

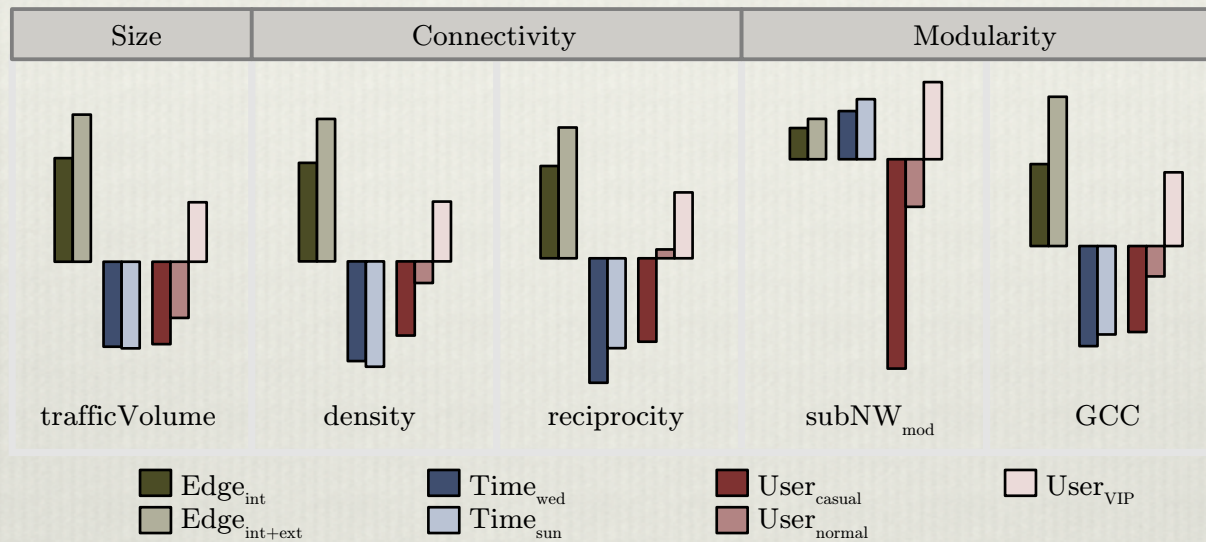


*z-score*



# NETWORK-LEVEL METRICS

## Edge-based networks

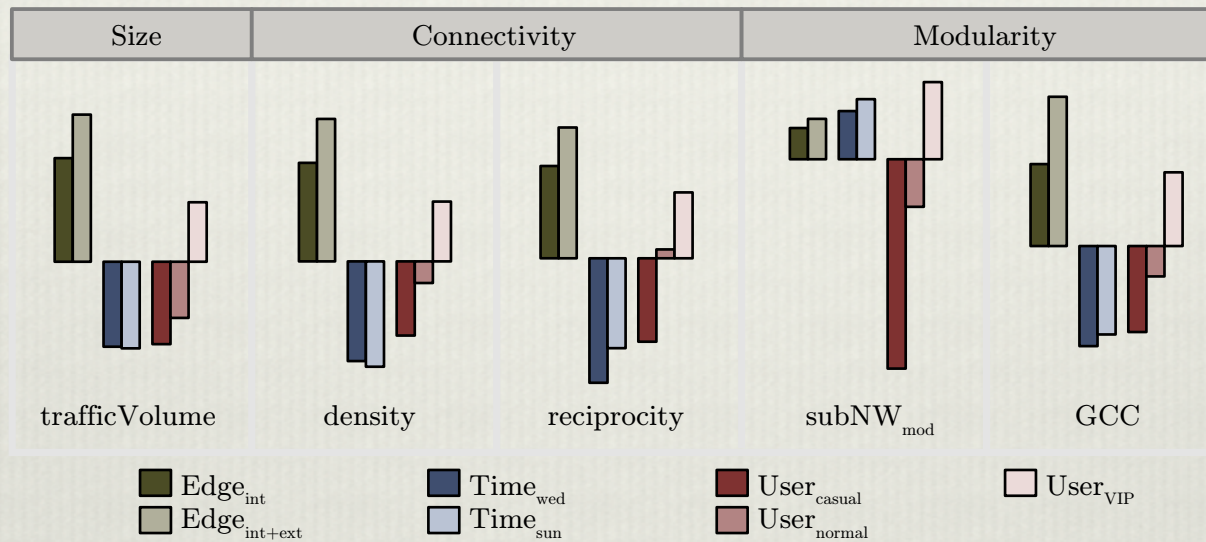


### Leaving the network does not necessarily entail less engagement

- Traffic increases significant when accounting for returning traffic [18.26%]
- Connectivity increases → User use new navigation paths
- Modularity increases → Users stay in the same part of the network

# NETWORK-LEVEL METRICS

## *Time-based networks*

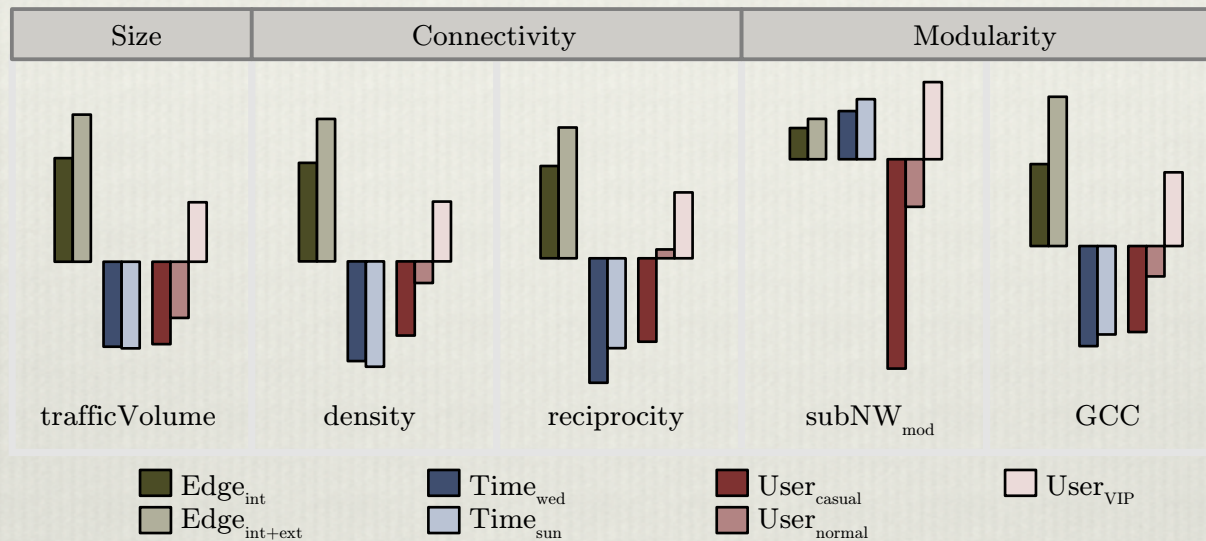


### **Navigation is more goal-oriented during the week**

- TrafficVolume and density are lower on Sunday → Less activity
- Reciprocity, subNW, and GCC are higher on Sunday → User return more often to already visited sites and visit more sites than usual (browsing with less specific goals)

# NETWORK-LEVEL METRICS

## User-based networks



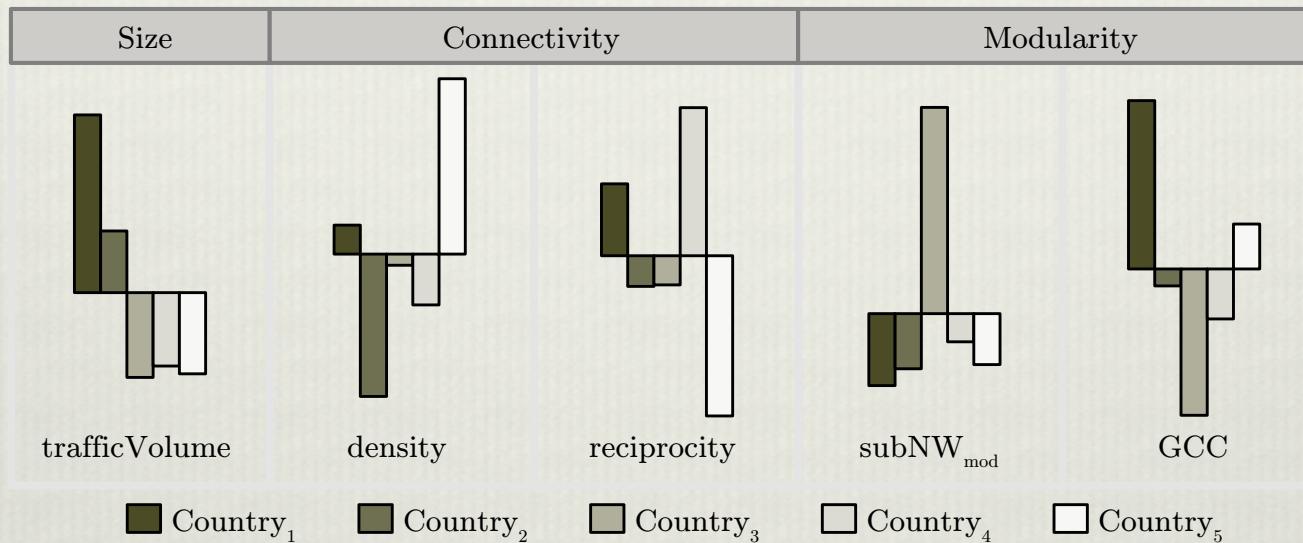
### Loyal users have also a higher networked user engagement

- TrafficVolume is higher for VIP user → More activity
- Connectivity and modularity is higher for VIP user → Users are interested in more sites
- GCC is lower for casual users → Users access sites using the front pages



# NETWORK-LEVEL METRICS

## Country-based networks



### Networked user engagement differs between countries

- Country<sub>1</sub>: Highest trafficVolume + GCC, high density + reciprocity, lowest subNW<sub>mod</sub> → high engagement
- Country<sub>5</sub>: Low trafficVolume, but high density + GCC, low subNW<sub>mod</sub> → less popular, but high engagement

# CONCLUSION AND FUTURE WORK

- Network analysis enhances the understanding of user engagement. It allows:
  - To analyze users browsing behavior on a global scale
  - To compare networks of sites (e.g. of different countries)
  - To observe differences over time
- Leaving the network does not entail less engagement
- Some network metrics are more useful than others

## Future work:

- Definition of metrics that combine network traffic and engagement
- Comparing traffic- and hyperlink-networks
- Studying the effect of changes in the network



**Janette Lehmann**  
Universitat Pompeu Fabra, Spain  
lehmannj@acm.org

**Mounia Lalmas**  
Yahoo Labs London  
mounia@acm.org

**Ricardo Baeza-Yates**  
Yahoo Labs Barcelona  
rbaeza@acm.org

**Elad Yom-Tov**  
Microsoft Research, Herzliya, Israel  
elad@ieee.org