A Method to Evaluate the Synergic Effect in Collaborative Information Seeking

Roberto González-Ibáñez
Chirag Shah
Muge Haseki

2011
CIS: SYNERGY

Is 1+1 > 2 in CIS?

If so... why, how, when this happens?

“The whole greater than the combination of the individual”

Buckminster Fuller, R. (1975)
BACKGROUND

• Users’ behaviors (Hyldegard, 2006; 2009)

• Comparison of CIS and (Joho et al.)

• Little is known about the synergic effect in CIS
  • Lack of methods/measures for evaluating synergy
    • Precision and Recall are not enough
A Method to Evaluate Synergy

• Three main components:
  • A procedure to experimentally evaluate synergy
  • A set of measures for evaluation
  • A system for experimentation
**METHOD: PROCEDURE**

- Experimental Design
  - 1+1 > 2
  - At least two experimental conditions
    - Single users => *Psudocollaboration* (baseline)
      - All possible unique combinations
    - Teams => Real collaboration
      - Common ground among participants
        - Friends
        - Couples
**METHOD: EVALUATION**

- Evaluation Measures
  - Universe:

  \[
  U = \bigcup_{i} \text{Coverage}(t) \quad \ldots \quad (1)
  \]

  - Relevant pages:

  \[
  U_r = \bigcup_{t} \text{ RelevantCoverage}(t) \quad \ldots \quad (2)
  \]
**Method: Evaluation**

- Evaluation Measures
  - Precision, Recall, F-Measure:

\[
Precision(t) = \frac{RelevantCoverage(t)}{Coverage(t)} \quad \ldots (3)
\]

\[
Recall(t) = \frac{RelevantCoverage(t)}{U_r} \quad \ldots (4)
\]

\[
F = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad \ldots (5)
\]
Method: Evaluation

• Evaluation Measures
  • Coverage

\[ \text{Coverage}(t) = \{wp_i : wp_i \text{ was visited by } t \land wp_i \in U \} \quad \ldots \ (6) \]

• Unique Coverage

\[ \text{UniqueCoverage}(t) = \text{Coverage}(t) \setminus \bigcup_{t_j \in (T \setminus \{t\})} \text{Coverage}(t_j) \quad \ldots \ (7) \]
METHOD: EVALUATION

- Evaluation Measures
  - Relevant Coverage

\[
\text{RelevantCoverage}(t) = \text{Coverage}(t) \cap U_r \quad ... \quad (8)
\]

- Unique Relevant Coverage

\[
\text{UniqueRelevantCoverage}(t) = \text{UniqueCoverage}(t) \cap U_r \quad ... \quad (9)
\]
**Method: Evaluation**

- Evaluation Measures
  - Useful Webpages: dwell time
  - Likelihood of discovery: effectiveness
    - Exploration/Diversity
  - Query diversity
**Method: Evaluation**

- Behavioral/User Measures
  - Cognitive Load => NASA’s TLX (Task Load Index)
  - Emotional Experience => PANAS, dynamic measures
  - Engagement
Information Retrieval

UNIVERSE

RELEVANT UNIVERSE

Likelihood of Discovery \( (t) \)

Coverage\( (t) \)

Useful WebPages\( (t) \)

Unique Useful WebPages\( (t) \)

Unique Coverage \( (t) \)

F-Measure\( (t) \)

Recall\( (t) \)

Precision\( (t) \)

Unique Relevant Coverage \( (t) \)

Query Diversity
Lavenshtein distance

User Measures
- Cognitive Load
  - NASA’s Task Load Index (TLX)
- Engagement
METHOD: EXPERIMENTAL SYSTEM

Coagmento

Chat System

Users'/Teams' Resources
APPLICATION

• Large user study:
  • 80 pairs working in eight collaborative conditions
  • 10 single users => 245 artificial teams
C1: Single Users  
- 10 users -

C2: Artificial Teams  
- 245 artificial pairs -

C3: Co-located  
(Same Computer)  
- 10 pairs -

C4: Co-located  
(Diff. Computer)  
- 10 pairs -

C5: Remotely Located  
- 10 pairs -
**APPLICATION**

- Published results:
  - Single users, co-located same computer, co-located different computers, remotely-located with chat communication
  - Two people collaborating is not the same as single users
    - Better results in coverage and diversity for real teams
    - Similar cognitive load for some conditions
    - Affective dimension positive for those working in collaboration
CONCLUSIONS

• Collaboration often useful for solving complex problems
  • Cost/Benefits
• Difficult to measure if and how collaboration will pay off
• IR measures are not completely useful in this domain to evaluate synergy.
• Our proposal provides a way to empirically evaluate various aspects of CIS
• We have successfully applied this method in a large user study
  • Results so far inform that 1+1>2
• Data collected from setups like those used in our study may help to explain the synergic effect in CIS
• This method as well as measures could be applied in other setups
CONCLUSIONS