



«Where should we fix this bug? a two-phase recommendation model»

2013

TSE

«A topic-based approach for narrowing the search space of buggy files from a bug report»

2011

ASE

«Compositional vector space models for improved bug localization»

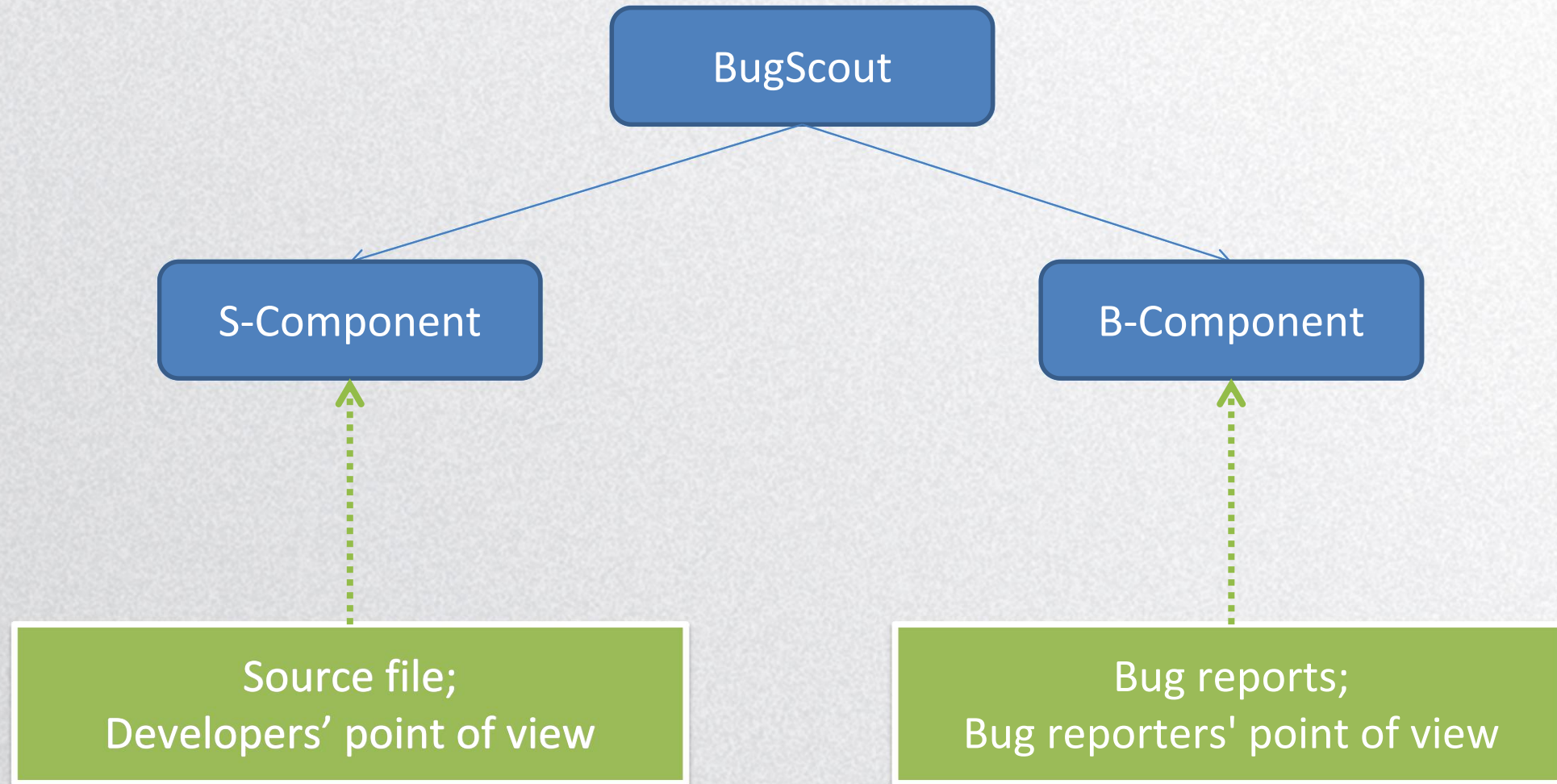
2014

ICSME



A topic-based approach for narrowing the search space of buggy files from a bug report

2011 ASE



A topic-based approach for narrowing the search space of buggy files from a bug report

2011 ASE

Topic Proportion

Topic Vector

Vocabulary and Word
Selection
(Distribution)

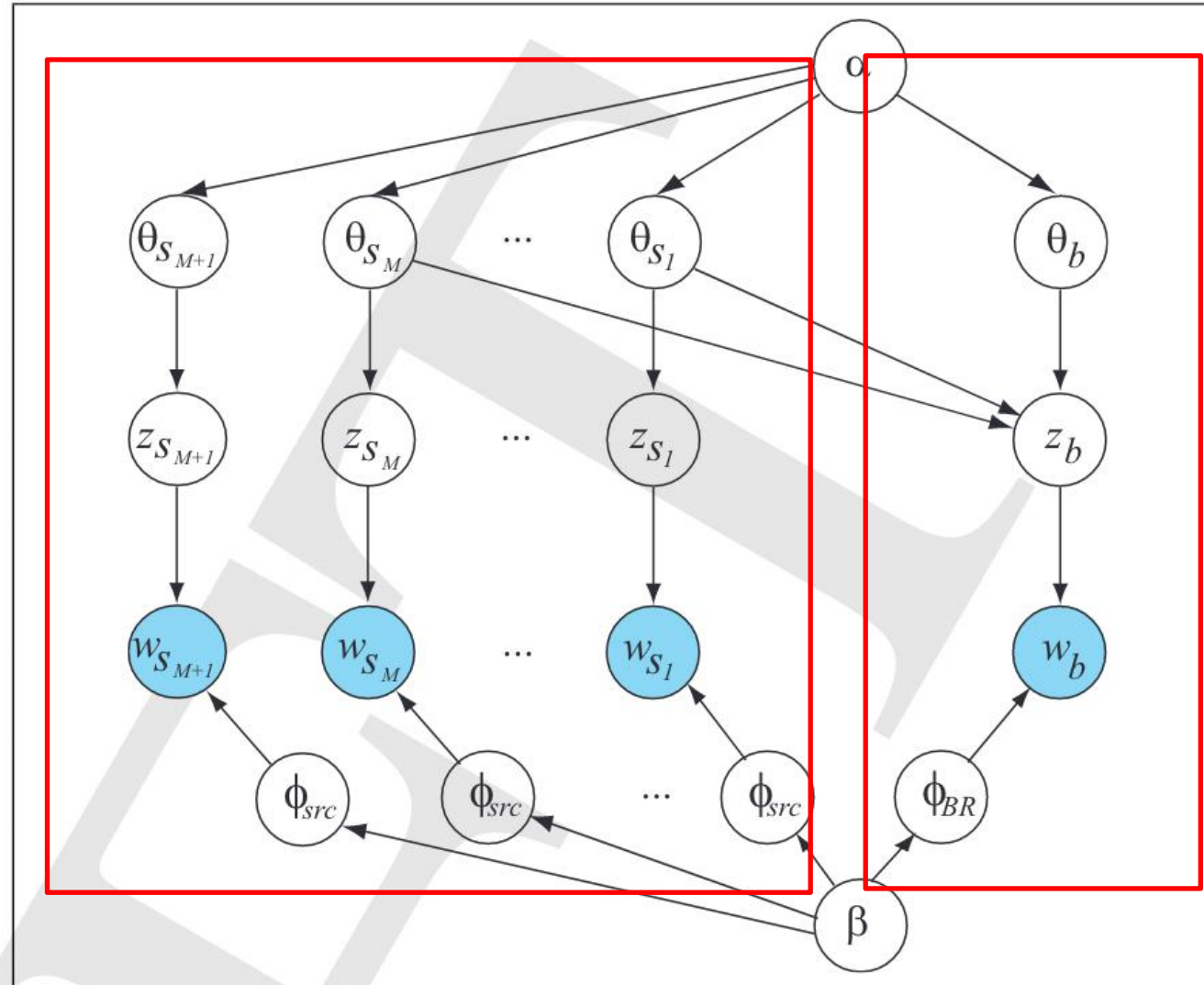


Fig. 6: BugScout Model



Where should we fix this bug? a two-phase recommendation model

2013 TSE

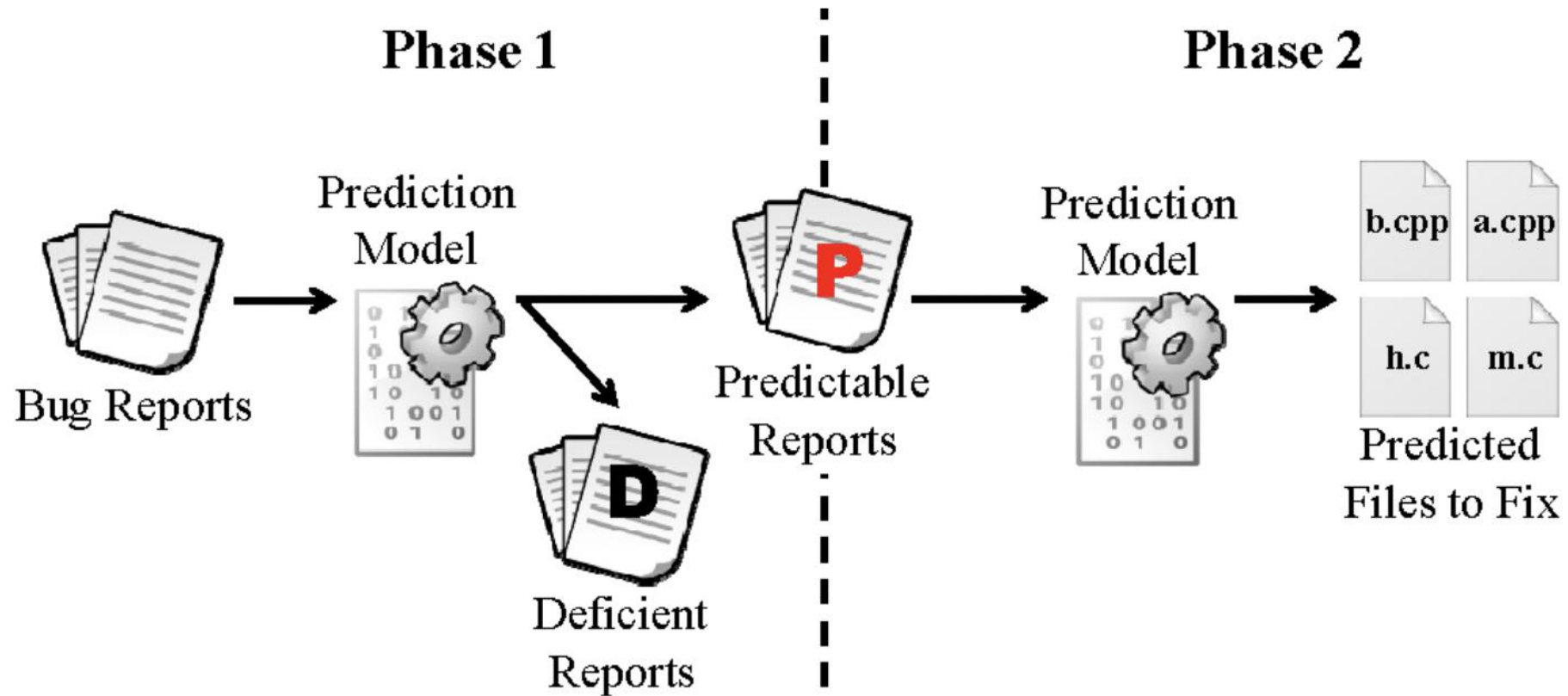
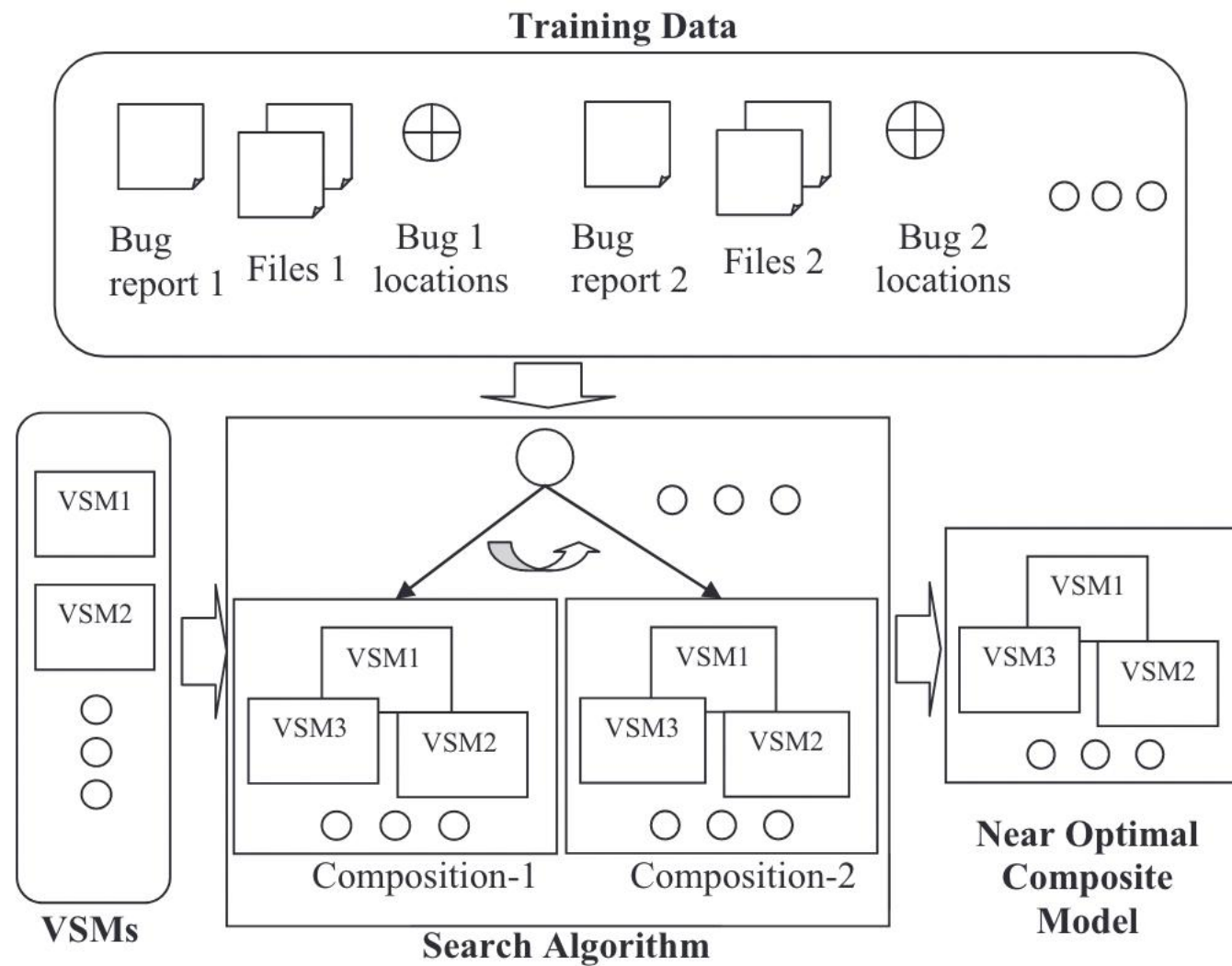


Fig. 3. The two-phase prediction model. This model recommends files to fix only when the given bug report is determined to have sufficient information.

Compositional vector space models for improved bug localization

2014 ICSME



(a) Training Phase

Compositional vector space models for improved bug localization

2014 ICSME

TABLE II. VARIANTS OF TF AND IDF

Term frequency	
$tf_n(t, d)$ (natural)	$ \{t t \in d\} $
$tf_l(t, d)$ (logarithm)	$1 + \log(tf_n)$
$tf_L(t, d)$ (Log ave)	$\frac{1 + \log(tf_n(t, d))}{1 + \log(ave_{t \in d}(tf_n(t, d)))}$
$tf_a(t, d)$ (augmented)	$0.5 + \frac{0.5 \times tf_n(t, d)}{max_t(tf_n(t, d))}$
$tf_b(t, d)$ (boolean)	$\begin{cases} 1 & \text{if } tf_n(t, d) > 0 \\ 0 & \text{otherwise} \end{cases}$
Document frequency	
$idf_n(t, D)$ (no)	1
$idf_l(t, D)$ (standard)	$\log \frac{ D }{df_t}$
$idf_r(t, D)$ (ratio)	$\max\{0, \log \frac{ D - df_t}{df_t}\}$

Name	Equation
tf_n-idf_n	$tf_n(t, d) \times idf_n(t, D)$
tf_n-idf_l	$tf_n(t, d) \times idf_l(t, D)$
tf_n-idf_r	$tf_n(t, d) \times idf_r(t, D)$
tf_l-idf_n	$tf_l(t, d) \times idf_n(t, D)$
tf_l-idf_l	$tf_l(t, d) \times idf_l(t, D)$
tf_l-idf_r	$tf_l(t, d) \times idf_r(t, D)$
tf_L-idf_n	$tf_L(t, d) \times idf_n(t, D)$
tf_L-idf_l	$tf_L(t, d) \times idf_l(t, D)$
tf_L-idf_r	$tf_L(t, d) \times idf_r(t, D)$
tf_a-idf_n	$tf_a(t, d) \times idf_n(t, D)$
tf_a-idf_l	$tf_a(t, d) \times idf_l(t, D)$
tf_a-idf_r	$tf_a(t, d) \times idf_r(t, D)$
tf_b-idf_n	$tf_b(t, d) \times idf_n(t, D)$
tf_b-idf_l	$tf_b(t, d) \times idf_l(t, D)$
tf_b-idf_r	$tf_b(t, d) \times idf_r(t, D)$



Compositional vector space models for improved bug localization

2014 ICSME

