Lexical Modeling of *Yamabuki*, Japanese Kerria in Classical Japanese Poetry

Hilofumi Yamamoto / Tokyo Institute of Technology
yamagen@ryu.titech.ac.jp

**Introduction**

- We conduct a lexical study of classical Japanese poetry using network modeling.
- The terms *yamabuki* (kerria), *kahazu* (frog), and *ide* (placename) are contained in some poetic dictionaries as entry items or collocations, and we have confirmed that they have strong relationships with each other.
- We have discovered the hub node term *yahe* in network models. The term *yahe* is, however, not recorded in any poetic dictionaries even as a single term.

**Material: Hachidaishū**
The eight anthologies compiled by the order of Emperors (ca. 905–1205), which contains about 9,500 poems.

**Calculation methods:**

\[ w(t, d) = (1 + \log \text{tf}(t, d)) \cdot \text{idf}(t) \]
\[ cw(t_1, t_2, d) = (1 + \log \text{ctf}(t_1, t_2, d)) \cdot \text{cidf}(t_1, t_2) \]
\[ \text{cidf}(t_1, t_2) = \sqrt{\text{idf}(t_1) \cdot \text{idf}(t_2)} \]
\[ \text{idf}(t) = \log \frac{N}{df(t)} \]

**Result**

Figure 2: Graph model of *kahazu* (蛙, frog) before pruning node 蛙.

Figure 3: Graph model of *kahazu* (蛙, frog) after pruning node 蛙.

Figure 4: Graph model of *Yamabuki*: a core node, 山吹 *yamabuki*, is pruned. *kahazu* (蛙, frog), *ide* (井手, place name, proper name), and *yahe* (八重, eightfold or double flower) are observed as hub nodes.

A minor term *yahe* (eightfold) can be shown as a hub node which plays a major role in connecting a topic word with other peripheral words which support/demonstrate poem stories. These minor words are not seen in poetic term dictionaries.

**Conclusion**

1. Discern not only patterns described by experts but also patterns yet undescribed, and
2. Identify not only specific or tangible words but also abstract or conceptual words which have a tendency to be left out of dictionaries.