

As part of our attempt to automatically classify users by their gang membership, we tried a variety of different features. One of the features that was helpful was a measure of how often a user mentioned a term for a gang. The intuition behind this is simple: someone who is a member of the Crips gang will talk about the Crips more often. The measure is simply how often a reference to a gang occurs per post. We also included features for gang-specific insults, with the idea that if someone uses insults for Crips often, they are likely not a Crip. The last type of reference is words that refer to gang alliances like the People Nation and Folks Nation. Members of those Chicago-based gangs frequently refer to their gang as the “Almighty [gang name] Nation”.

We used regular expressions to capture occurrences of these words and variations on them such as the use of the orthographic substitutions mentioned previously, plurals, feminine forms, etc. Additionally, in the Blood and Hoover features, they sometimes use numbers to replace the ‘o’s representing the street that their gang is located on. So the Bloods from 34th Street, say, might write “B134d”.

The following table shows some of the types of words that would be included in each category.

| Feature | Basic forms covered |
|----------------------------|----------------------------------|
| Crip mentions | crip, loc |
| Crip insults | crab, ckrip, ck |
| Blood mentions | blood, damu, piru, ubn |
| Blood insults | slob, bklood, pkiru, bk, pk |
| Hoover mentions | hoover, groover, crim, hgc, hcg |
| Hoover insults | snoover, hkoover, hk |
| Gangster Disciple mentions | GD, GDN, Gangster Disciple |
| Gangster Disciple insults | gk, dk, nigka |
| Folk mentions | folk, folknation |
| People mentions | people, peoplenation |
| Nation usage | almighty, nation |
| Latin King mentions | alkqn, king, queen |
| Black P. Stone mentions | stone, abpsn, moe, black p. |
| Vice Lord mentions | vice, lord, vl, avln, foe, 4ch |
| Trinitario mentions | trini, trinitario, 3ni, 3nitario |

Crip: (? :crip|cripp?ing?|locc?(ette)?) [sz]?
Crab: (? :ck|ckr(ab|ip)|crab) [sz5]?
Blood: (? :bl(? :xx|u)d|blood[sz5]?|b[l11]\d\d\d?d|damu|piru|ubn) [sz5]?
Slob: (bk|pk|bkl(? :oo|xx)d|pkiru|slobk?) [sz]?
Hoover: ((? :h|gr)(? :oo|xx|\d\d\d?)v(? :a|er?)|crim|hgc|hcg) [sz]?
Snoova: (? :hk|(? :hk|sn)(? :oo|xx)v(? :a|er)) [sz]?
GD: (? :[bim]?gd|a?b?gdn|(? :black|insane|mania?c)[]?gangst(? :a|er)|dis?c?iple) [sz]?
GDk: (? :gk|g?dk|nigk(? :a|uh)) [sz]?
Folk: (? :folkk?|folknation) [sz]?
People: (? :people|peoplentation) [sz5]?
Nation: (? :(? :al)?mighty|nation)
Latin King: (? :a?lkq?n?|king|queen) [sz5]?
Stone: (? :[s5]tone|a?bp[s5]n?|black p\.|moe|moor) [sz5]?
Vice Lord: (? :vicelord|vice|lord|a?[c4]?vln?|foe|4ch) [sz5]?
Trinitario: (? :(? :3|tri)ni(? :tario)?) [sz]?